>TrainDL

>Teacher training for Data Literacy & Computer Science competences // Report Round 1 // **Deliverable 4.4: Report on First Evaluation Phase**

train-dl.eu

Martin Reger; Evgenia Samoilova; Ulrike Lucke; Anastasia Tsymboulova

University of Potsdam

Content

List of Figures

List of Tables

Executive Summary

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. The project applies an action research approach to design, deliver and evaluate DL and AI teacher trainings in three countries: Austria, Germany, and Lithuania. The project comprises three intervention rounds targeting computer science teachers as well as teachers from other subjects at the primary and secondary levels. During the first round of interventions (2021-2023) focusing on computer science (CS) teachers, a comprehensive review of European educational policies was undertaken to assess the integration of DL and AI into teacher education. Based on this analysis, as well as additional feedback from stakeholders in the three countries, the TrainDL team developed and carried out teacher trainings for pre- and in-service CS teachers at the secondary level. This report includes the evaluation results of the four trainings within the first round. The evaluated trainings included three 7-hour sessions conducted between June 2022 and January 2023 in Berlin, Germany; Vilnius, Lithuania; and Vienna, Austria. Additionally, in November 2022 a 3-hour training session, split over two days, was held specifically for pre-service CS teachers in Berlin, Germany. Within the trainings, pre- and in-service CS teachers were introduced to such topics as machine learning, classical AI, and data lifecycle.

The evaluation of training sessions mainly focused on in-service CS teachers and their ability to integrate DL and AI into their classrooms post-training. Additionally, the evaluation included the pre-service CS-teacher session aimed at fostering positive attitudes towards DL and AI. Training sessions were evaluated using pre- and post-training surveys, DL and AI self-assessment and knowledge tests, and semi-structured interviews conducted immediately and approximately six months post-

training. We used a mixed methods approach, combining quantitative data on participants' characteristics and competences with qualitative insights from teacher interviews. The qualitative interviews primarily focus on the overall training feedback and experiences of, as well as barriers to integration of DL and AI into the classroom. The evaluation for the 3-hour pre-service training was simplified, consisting solely of the post-training survey.

The quantitative results of the evaluation showed, that the trainings in all three countries had some positive impact on teachers' perceived competences on how to use AI content in class, as well as teachers' understanding of AI concepts introduced in the training. For the DL content, mostly composed of the data lifecycle topic and the Orange3 tool, the results were mixed: improvements in this area demonstrated high variability with no evident increase post-training. On average, compared to DL (specifically the topic of data lifecycle using the Orange3 tool), AI topics and exercises were perceived to be more suitable for teaching in the classroom.

Participants in all three countries tend to agree, that DL and AI content is missing in the current CS curriculum. Furthermore, they agree on the societal importance of the DL and AI topics and expect these subjects to generate substantial student interest and engagement. However, there is not enough evidence to argue that a single 7-hour training session is sufficient for the integration of DL and AI topics into teaching as measured immediately after the training. The follow-up data from Germany, collected ca. 6 months after the training and limited to a very small sample size of seven, suggest that the integration of DL and AI topics can be challenging for those who did not teach these topics prior to the TrainDL training. The follow-up data-collection and analysis in Lithuania and Austria is ongoing and will be described in the final report.

The qualitative findings across the three countries demonstrated that teachers recognize the urgency to integrate DL and AI into teaching and curriculum framework, given their societal, political, and practical relevance. However, the primary step for

such integration is training teachers to gain sufficient proficiency in these two areas, DL and AI.

Participants in all three countries reported that the TrainDL training served as a valuable introduction to AI and DL. However, the depth and complexity of certain topics, especially in the realm of AI, were challenging for participants. Exercises involving the Orange3 tool consistently received mixed feedback. While some found it suitable for teaching, others saw it to be rather complex, another critique involved its graphical interface and lack of programming as deterrents for more advanced students.

In all three countries, participants highlighted a difficulty of integrating DL and AI into their teaching. While teachers felt more prepared to do so after the TrainDL training, the integration would require further engagement with the material. A single 7-hour course is not enough to ensure integration of DL and AI into teaching, especially if teachers did not teach these subjects prior to the training. Teachers expressed a strong desire for concrete course plans and materials tailored to specific grade levels to streamline integration into teaching and reduce their preparation workload.

In every country, teachers emphasize the urgent need to incorporate DL and AI into the CS framework curriculum. Some teachers argue for integrating DL and AI not only into CS classes but also into other non-CS subjects. In Germany and Austria, educators recommend making CS a compulsory subject in secondary education. In Lithuania, there's a suggestion to introduce CS already at the primary level (as it is a compulsory subject for the lower secondary level). In all the countries, teachers pointed out the challenges of revising the already content-heavy curriculums to make room for DL and AI, stressing the importance of reprioritising content. At the same time, there's a shared understanding that merely adding DL and AI to the curriculums won't ensure, that teachers are immediately adept at teaching them, underscoring the continuous need for teacher training.

Most teachers agree that a single 7-hour training is not sufficient to acquire sufficient competences to teach DL and AI in class. At the same time, some teachers have difficulties attending even a one-day training, due to their teaching responsibilities and high workload. Therefore, the challenge lies in designing a training program that is both comprehensive and flexible. Such a program could be modular and spread over multiple sessions, allowing teachers to grasp the complexity of DL and AI at their own pace, without compromising their teaching hours. It is also crucial to understand and address the underlying factors that motivate and incentivise teachers to prioritise such training amidst their already demanding roles.

Teachers also voiced concerns about readily accessible Large Language Models (LLMs), particularly regarding academic integrity. This underscores the need for comprehensive AI support, not only in terms of teaching methodologies but also in regular teaching practices (e.g., assessment).

For the only pre-service training in Berlin, the quantitative post-survey results show overwhelmingly positive attitudes towards incorporating DL and AI into future teaching among the surveyed pre-service teachers.

The overall positive feedback of the TrainDL training as a starting point of introduction to DL and AI, coalesced with teachers' strong belief in the high relevance of DL and AI for teachers, students, and society at large, suggests a promising possibility for the integration of these subjects into the classroom. However, the identified limitations of this study, such as the small sample size and limited variation in the format, underline the need for more comprehensive research to refine and optimise such teacher training programs. Furthermore, the challenges brought to the foreground - such as for example, teacher workload, rapid technological advancements, and limited capacity of the framework curriculum - offer an opportunity to reflect on the design of future training and respective policies.

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 other subjects at the primary and secondary levels. During the first round of interventions focusing on CS teachers, a comprehensive review of European educational policies was undertaken to assess the integration of DL and AI into teacher education (see Deliverable 1.9). Based on this analysis as well as additional feedback from stakeholders in Germany, Lithuania, and Austria (see Deliverable 1.6) the TrainDL team developed teacher training for pre-service and in-service CS teachers at the secondary level (see Deliverable 2.3). This report presents the evaluation results of the first round of interventions, focusing on the implemented training concepts targeting pre- and in-service computer science (CS) teachers at the secondary level. The evaluated trainings included three 7-hour sessions conducted between June 2022 and January 2023 in Berlin, Germany; Vilnius, Lithuania; and Vienna, Austria. Additionally, in November 2022, a 3-hour training session, split over two days, was held specifically for pre-service CS teachers (university students) in Berlin, Germany. Within the trainings, pre- and in-service CS teachers were introduced to such topics as machine learning, classical AI, and data lifecycle (see Deliverable 2.3 for further details).

The evaluation of the first intervention round primarily focuses on teachers' capacity to integrate the acquired content on DL and AI into their teaching, as measured

immediately after the trainings. The follow-up data collection, conducted approximately six months after the training session, provides additional insights into the actual integration or lack thereof of these topics into teaching.[1]

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

- The evaluation survey administered before and immediately after the training, as well as six months after the respective training as a follow-up;
- 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 and approximately six months after the training (as a follow-up).

Table 1 presents an overview of the evaluated training sessions and the evaluation instruments used. Since the 3-hour training for the pre-service teachers in Berlin was not the main focus of the evaluation, it was evaluated only via the post-evaluation survey and pre- / post-self-assessment and knowledge test.

Date	Location	Target group	Duration	Evaluation instruments used	Number of participants
13.06.2022	Berlin, Germany	Pre- and in-service CS teachers at the	7 hours	• (Pre- and post-) evaluation survey • (Pre- and post-) DL and AI self-assessment and knowledge test	24

[1] While the follow-up data-collection and analysis for the CS training in Germany is completed, the analysis of the follow-up in Lithuania and Austria is still ongoing and will be incorporated into the final report (D4.7).

		secondary level		• Semi-structured personal interviews • Six-month follow-up: evaluation survey • Six-month follow-up: semi-structured online interviews	
02-03.11.2022	Berlin, Germany	Pre-service CS teachers at the secondary level	3 hours	• Post-evaluation survey	25
10.12.2022	Vilnius, Lithuania	In-service CS teachers at the secondary level	7 hours	• (Pre- and post-) evaluation survey • (Pre- and post-) DL and AI self-assessment and knowledge test • Semi-structured personal interviews	21
31.01.2023	Vienna, Austria	In-service CS teachers at the secondary level	7 hours	• (Pre- and post-) evaluation survey • (Pre- and post-) DL and AI self-assessment and knowledge test • Semi-structured personal interviews	25

Table 1 Overview of the evaluated trainings of the first intervention round: date, location, target group, duration, evaluation instruments used, number of participants

This report is structured into seven primary sections. The first section delineates our research questions and hypotheses, outlines the evaluation methodology for the trainings, and describes the data collection instruments used. The outcomes for each training are detailed separately in sections 2 through 5. Finally, sections 6 and 7 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. With the objective of gaining insights into the scarcity of teacher training programs on AI and DL and bridging this gap, the project opted for an action research methodology (Baskerville and Wood-Harper 1996; Burns 2010). Action research is characterized by its iterative nature, involving multiple rounds of planning, action, observation, feedback, and reflection. This report specifically addresses the research questions related to the first round of interventions designed for secondary CS pre- and in-service teachers. Though the evaluation of the training sessions primarily geared towards in-service teachers centred on their capacity to integrate DL and AI into their classrooms, an extra pre-service CS session was also assessed, with the goal of promoting positive attitudes towards DL and AI. This first intervention consists of the three steps:

Phase 1 - Understanding Practice: The project team conducted research on the integration of AI and DL in school education and explored the availability and common practices of teacher training for CS teachers. Furthermore, the project examined educational policies and engaged with policymakers and stakeholders from Germany, Austria, and Lithuania to identify relevant factors for teacher education in the context of AI and DL.

Phase 2 - Deliberate Improvements: Based on the findings from the first step, a teacher training concept for CS pre- and in-service teachers at the secondary level, was developed.

Phase 3 - Implementation and Observation of Improvements: The trainings were implemented in Germany, Austria, and Lithuania. To assess the effectiveness of the training sessions, we have evaluated them using methods and data described below.

To evaluate the three 7-hour CS 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 further illuminated the ability to integrate DL and AI into the teaching, and also examined relevant contextual factors and obstacles. The evaluation of the additional 3-hour pre-service training followed a more simplified procedure, omitting the pre-evaluation survey, qualitative interviews, and follow-up data collection: only a post survey was conducted.

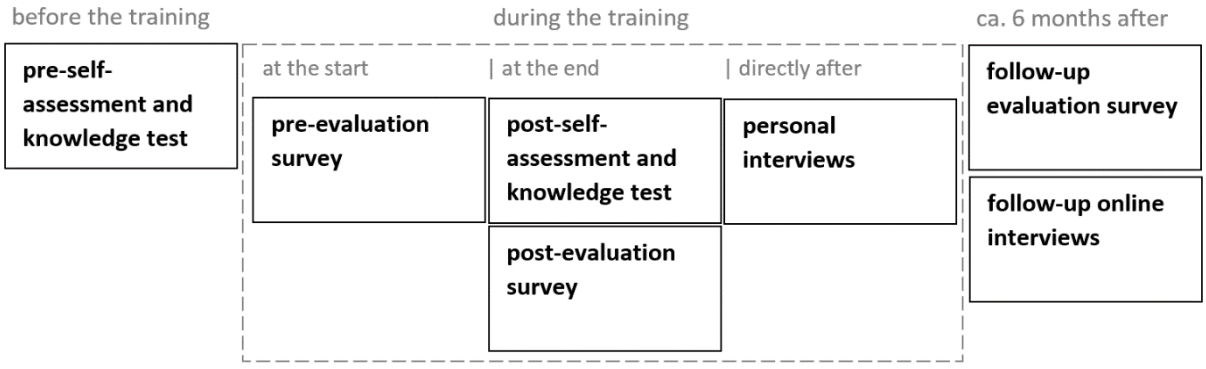


Figure 1 Overview of the evaluation process for each 7-hour training session

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 pseudony-

misation code (see Appendices 1-5 for translated versions of the surveys for Germany[2]), 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.

In the training, the focus of DL primarily revolved around the concept of the data lifecycle.

[2] Questionnaires and interview guides in Germany, Lithuania, and Austria were almost identical with exception of the country-specific questions (e.g., types of schools, states, subject names).

1.1 Quantitative data: research questions, instruments, and analysis

The quantitative data of the first intervention cycle was primarily used to address the hypothesis 1 and 4[3], derived from the two research questions:

(1) *What is the effect of the designed one-day DL and AI training on the ability of in-service CS teachers with a solid background in CS to integrate DL and AI into their classes?*
(2) *What is the effect of the designed lecture on DL and AI teaching methodologies on the attitudes of pre-service CS teachers towards integrating AI and DL into their future teaching?*

Table 2 describes the two guiding hypotheses used for the first intervention round. The first hypothesis for in-service CS trainings, deals with teachers' ability to incorporate DL and AI into their teaching. 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.

[3] Numbers of the hypotheses refer to the Deliverable 4.3, where each developed hypothesis was assigned a number.

Hypothesis # as assigned in Deliverable 4.3	Hypothesis text
Hypothesis # 1	If in-service CS teachers with a solid background in CS participate in a one-day DL and AI teacher training workshop, they then are able to integrate DL and AI in their CS classes.
Hypothesis # 4	If pre-service CS teachers participate in a specific lecture on DL and AI teaching methodologies, they then have positive attitudes towards integrating DL and AI into their future teaching.

Table 2 Guiding project hypothesis used for the first round of interventions for the pre- and in-service CS teachers in Berlin, Germany; Vilnius, Lithuania, and Vienna, Austria

For the pre-service CS training, we have looked at several post-survey items developed specifically for this training, with a focus on attitudes towards teaching DL and AI.

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

- teachers' expectations about students' interest and their perceived ability to generate student interest and engagement for the topics of DL;
- 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.

To collect quantitative data, we have used two instruments:

- the pre- and post-evaluation survey developed by the University of Potsdam, and
- the DL and AI self-assessment and knowledge test, developed by the Freie Universität Berlin

To analyse the pre- and post-data, we employ descriptive statistics and the Wilcoxon signed-rank test. This test uses mean ranks to assess, whether there is a statistically significant difference between two related samples: the pre- and post-measures, taken from the same individuals. If the Wilcoxon signed-rank test is statistically significant, this supports the conclusion that there is a difference between the pre- and post-measures. However, the test statistic and p-value from the Wilcoxon signed-rank test do not tell us the direction of the difference (i.e., which group has higher values). Therefore, to interpret the differences, we will look at the descriptive statistics with a focus on measures of variability for non-parametric data - median and interquartile range (IQR) that describes where the middle 50% of the data falls. To visualize the data we mostly use boxplots (that display the median, IQR, and possible outliers), which are very useful in comparing distributions between groups (i.e., pre- and post-measures as well as differences between the countries).

The benefit of using related samples (measures from the same individuals for pre- and post-data collection) is that it reduces variability caused by individual differences. As a result, it potentially lowers the sample size requirements. However, it's important to note that our samples are small, so results should be interpreted with caution. While small sample sizes notably limit the generalizability of the study, they provide indications of trends within the sample and can be valuable when combined with the qualitative results of the interviews. Despite the consistent format and content of the 7-hour training across the three countries, we opted not to merge the data

and analyse the results separately for each country. This was due to significant variances in the school systems, the integration levels of DL and AI, teacher education, and potential cultural differences among the targeted countries.

The evaluation survey:

The questionnaires (see Appendices 1-2) included information on demographics (e.g., sex and age), educational background, type of the school where in-service teachers are employed, teaching hours and subjects, experience with DL and AI as well as attitudes towards these topics, expectations and the perceived ability to generate student interest. Also included are: engagement for the topics of DL, teachers' feedback on the learned content and format of the training, and most importantly teachers' perceived competences on how to use DL and AI in class. The follow-up survey for the in-service teachers focused on the actual integration of the content into the classroom.

The perceived competences on how to use DL and AI in class were measured via the following pre- and post-survey items. For each item, 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"):

- [measured in the pre- and post-survey] "I know how to use content about DL in the classroom."
- [measured in the post-survey] "I know how to use content about AI in the classroom."

In the post-evaluation survey the following items were included to provide additional insights on the aspect of pedagogical content knowledge and potential for DL and AI integration:

- [measured in the post-survey] "After the training, I have gathered enough competences to teach the learned content in class."
- [measured in the post-survey] "I am willing to invest time and effort to incorporate AI into my teaching."

The post-evaluation survey included a series of questions to assess participants' reactions to the topics and materials/exercises covered in the training:

- [measured in the post-survey] “How suitable did you find the choice of the topics on introducing DL and AI for your teaching?” (1-not suitable at all – 6 very well suited):
  - Classical AI
  - ML (reinforcement, supervised, unsupervised learning)
  - Data lifecycle and fundamentals of statistical data analysis and integration
- [measured in the post-survey] How suitable did you find the practical examples from the workshop for your teaching? (1-not suitable at all – 6 very well suited)
  - AI or not? AI Bingo
  - Shopping cart game
  - Beat the robot
  - The good monkey-bad monkey exercise
  - Project work on data analysis and interpretation using Orange3

For the hypothesis 4 targeting the pre-service CS training, the following items on attitudes towards DL and AI were included into the post-evaluation survey:

- [measured in post-survey] "After the lecture and seminar on November 2nd and 3rd, here is my attitude towards handling DL and AI in my future teaching: (1) - very negative, (2) – negative, (3) - slightly negative, (4) – neutral, (5) - slightly positive, (6) – positive, (7) - very positive.”

- [measured in the post-survey] "Do you want to integrate DL and AI into your future teaching? (1) – not at all - (6) – definitely"

The web-based surveys were programmed in QUAMP survey software (versions 3.4.07 – 4.4.3). Table 3 provides an overview of the response rates for the pre-, post-, and the follow-up versions of the evaluation survey for each training sessions:

Training	Number of participants	Number of participants completed the survey (%)			Number of participants completed the follow-up survey (about 6 months after the training)
		Pre	Post	Both	
13.06.2022, Berlin, pre- and in-service CS teachers	24	23 (96%)	21 (88%)	20 (83%)	7 (30%)
02-03.11.2022, Berlin, pre-service CS teachers	25	-	18 (72%)	-	Follow-up survey was not planned
10.12.2022, Vilnius, in-service CS teachers	21	21 (100%)	17 (81%)	16 (76%)	Follow-up survey was scheduled
31.01.2023, Vienna, in-service CS teachers	25	19 (76%)	15 (60%)	14 (56%)	7 (28 %)

Table 3 Response rates for the evaluation survey for each of the trainings: number and % of participants, who completed the test prior to the training, after the training, and both the pre- and post-versions

The DL and AI self-assessment and knowledge test:

The AI and DL self-assessment and knowledge test includes 25 self-assessment questions (15 for DL dimension and ten for AI dimension with a scale from "1" – strongly disagree to "5" – strongly agree) and 14 objective knowledge questions (four for DL dimension and ten for AI dimension) on understanding of the DL and AI concepts introduced in the training. Each question in the objective knowledge test carried a potential score ranging from 0 to 1. Participants had the opportunity to select multiple answers, with scores being deducted for selecting incorrect options. This test was administered before and directly after the training. Table 4 summarises response rates for the pre- and post-test for each of the trainings.

Training	Number of participants	Number of participants completed the test (%)		Number of participants completed both pre- and post-tests
		Pre	Post	
13.06.2022, Berlin, pre- and in-service CS teachers	24	19 (80%)	20 (83%)	16 (67%)
02-03.11.2022, Berlin, pre-service CS teachers	-	-	-	-
10.12.2022, Vilnius, in-service teachers CS	21	19 (90%)	18 (86%)	14 (67%)
31.01.2023, Vienna, in-service teachers CS	25	21 (84 %)	11 (44 %)	11 (44%)

Table 4 Response rates for the AI and DL self-assessment and knowledge test for each of the trainings: number and % of participants, who completed the test prior to the training, after the training, and both the pre- and post-test versions.

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?*
(5) *What potential changes could arise from integrating DL and AI into the framework curriculum and the teaching of these subjects?*
(6) *(follow-up) In retrospect, what personal and teaching insights have participants derived from the training?*
(7) *(follow-up) To what extent have participants incorporated DL and AI into classroom instruction after the training? If not, is there an intention to integrate them in the future?*
(8) *(follow-up) What areas, if any, do participants believe could be enhanced in both the training and the content on DL and AI?*

The in-service CS trainings were followed up by the two rounds of qualitative interviews. The interviews were conducted right after the training and (for the CS in-service training) about six months after the training. In both instances, a semi-structured interview guide developed by the University of Potsdam was used. The interviews directly after training were conducted in person, whereas the follow-up interviews were conducted online.

The interview guide:

The questions in the interview guide administered right after the training (see Appendix 5) focused on teachers' perception of the respective workshop, experiences in

integrating the DL and AI content into their classroom and barriers for such integration. Teachers were also asked about the importance of both topics for teacher training and for framework curricula, as well as their wishes for policymakers. The follow-up interview guide roughly six months after the training (see Appendix 6) is primarily focused on the integration of the training content into the classroom. Table 5 includes an overview of the number of interviewed participants for each training.

Training	Number of participants	Number of participants who took part in the qualitative interviews (%)	Number of participants participated in the follow-up interview (about 6 months after the training)
13.06.2022, Berlin, pre- and in-service CS teachers	24	6 (25 %)	2 (8%)
02-03.11.2022, Berlin, pre-service CS teachers	25	No interviews were planned	Follow-up survey was not planned
10.12.2022, Vilnius, in-service CS teachers	21	8 (38%)	Follow-up interview was scheduled
31.01.2023, Vienna, in-service CS teachers	25	6 (24%)	2 (8%)

Table 5 Number and % of participants, who took part in the qualitative interviews right after and ca. 6 months after the training.

While the qualitative interviews in Germany and Austria were conducted in German, the interviews in Lithuania were conducted in English. All the interviews were transcribed and analysed.

For the first training on 13.06.22 in Berlin, the interviews were analysed with the help of qualitative content analysis (specifically summarizing content analysis) (Mayring 2010). For the subsequent trainings, due to time constraints, we have opted for a more condensed version of the content analysis using the focused interview analysis approach (Kuckartz and Rädiker 2020). While the analysis according to Mayring (2010) is relatively strict and includes multiple work steps for all text segments - including paraphrasing, generalizing, and reducing - the focused analysis allows for greater flexibility in selecting which steps to shorten. For instance, the user can decide whether to summarize text segments or omit certain steps altogether. In both methods, we have used both deductive and inductive coding. While the former codes were developed based on the interview guide and applied to all the interviews, within them, an inductive code captured new information that emerged directly from the data. For all personal interviews, we analysed participants' familiarity (consisting of prior knowledge and previous experience) with the topics of DL and AI using the standardized scale approach developed by Maying (2010), employing a 3-point scale (no familiarity, moderate/average familiarity, high familiarity).

The qualitative results of the analysis in this report are presented according to the structure of the interview guide (see Appendix 5). The primary questions from the interview guide serve as the main themes under which inductively generated categories are presented. The related sub-questions align with the respective sub-themes.

Themes addressed in the qualitative analysis:

Below is a description of the main themes derived deductively from our research questions and aspects of the interview guide. These themes were developed to provide a structured framework for analysing the data in line with our key inquiry areas. Additionally, the analysis also led to the inductive emergence of specific categories, which were not predefined but surfaced naturally during the review of our qualitative data. These categories, revealing deeper insights into participant experiences and perceptions, will be detailed in the respective chapters for each training session.

A. Themes in interviews immediately after the training:

- Training: In this theme, some fundamental aspects related to the training are clarified. Firstly, we explore the participants' expectations and the factors that led to their participation in the training. Secondly, we delve into their personal perspectives on the difficulty level of the topics. Furthermore, we examine the alignment between the training content and participants' prior familiarity with the concepts of DL and AI, which includes their previous knowledge and experience.

- Teaching DL and AI & difficulties conveying DL and AI concepts: In this theme, our focus lies on exploring the integration of DL and AI topics into teaching practices prior and after the training. When it comes to the aspect of integration after the training, our objective is to assess participants' readiness and confidence in effectively incorporating DL and AI into their teaching. Furthermore, our interest extends to evaluating the challenges associated with conveying knowledge about DL and AI to students.

- Establishment and steps to integrate DL and AI: This theme is dedicated to exploring the integration of DL and AI topics into teaching, specifically within the context of school classrooms. It encompasses the following key aspects: participants' perspectives on the integration of DL and AI topics within the framework curriculum; identification of effective steps or strategies for embedding both topics into classroom and school curriculum; teachers' opinions regarding the integration of DL and AI topics in teacher education programs; lastly the exploration of any anticipated barriers that may hinder the successful integration of DL and AI.

- (Possible) changes through the integration of DL and AI in the framework curricula:
  This theme centres on potential changes resulting from the incorporation of DL and AI into framework curricula. Specifically, we query teachers about anticipated changes concerning students, school authorities, and other school-related aspects, as well as the broader societal implications.

- Training feedback and potential for improvement: This theme primarily focuses on training feedback. Specifically, we focus on participants' feedback on the length, content (topics and exercises), format (in-person event), and participant interaction (such as the balance between frontal and interactive parts). Alongside positive feedback, we are particularly interested in suggestions or criticisms that can be used to inform improvements for future trainings.

- Wishes for education policy: This theme describes teachers' wishes or suggestions for education policies concerning the topic of DL and AI in school education.

- Other themes: In this theme, we focus on addressing all other insightful statements that have the potential to enrich the context of DL and AI topics.

B. Themes for follow-up interviews:

- Training: This theme focuses on how the training is retrospectively perceived with some time elapsed and to what extent (if at all) the participants have benefited from it.

- Integration after the training: The theme describes if there has been any integration of DL or AI into teaching CS after the training. In addition, in the case of integration, we are interested in the experiences during the process and

details such as the duration of implementation, grade level, topics covered, etc. In the case of non-integration, we are interested in the reasons behind it or what would facilitate future integration.

- Training feedback and potential for improvement: With this theme, we identify possible improvements that we could implement in the future. Additionally, we are interested in general suggestions related to the DL and AI in the context of teacher training.

2. CS in-service training, Germany

2.1 Sample

Out of 24 participants, all of whom were from Berlin, 23 participants took part in the pre-evaluation survey. Figure 2 describes socio-demographic data of the participants as reported in the pre-survey. The average age of the participants in the training was 45 years old, ranging from 30 to a maximum of 63 years. The share of women constituted ca. 22% (five out of 23 participants). One person has chosen a non-binary category, "diverse." While 13 participants reported being fully trained teachers or having accomplished a state exam, five participants reported being career changers (entered teaching outside the traditional career path) and five were pre-service teachers at the final practical stage of their training ("Referendariat"). Most teachers indicated that they teach students ranging from seventh grade to twelfth grade. Three participants indicated teaching grades five and six. Half of the participants were employed in a gymnasium, while the other half reported other types of secondary schools.

The socio-demographic data of the participants is consistent with the 2022 data on CS teaching in Germany. According to the Informatik-Monitor by the German Informatics Society (Schwarz, Hellmig, and Friedrich 2022), Berlin has implemented a standardized framework curriculum for the optional course of computer science (CS)

across all school types, from grades seven to ten. Additionally, a curriculum for computer science is provided for the introduction phase (typically grade 11) and the qualification phase (usually grades 12-13) of the upper level of the secondary school ("Gymnasium"). Comparing the socio-demographic data of the participants to that of pre- and in-service CS teachers in Germany, especially in Berlin, is difficult. According to a report by Stifterverband and Heinz Nixdorf Foundation (Schröder, Suessenbach, and Winde 2022), although such data is partly available, they are scattered across ministries and statistical offices of the federal states and are often available only upon request. Schröder et al. (2022) report that depending on the federal state, the percentage of female computer science teachers ranges from 24 to 45%. The authors of the report state that the share of women among CS students is also notably low: three out of ten first-year students, one out of four current students, and one out of four graduates in the field of CS education are female (2022, 12).

In terms of age group distribution for secondary school teachers, similar to the Berlin sample of the CS teachers, the data provided by the OECD for all the teachers at the secondary level are fairly evenly spread: the 30-39 age bracket accounted for 28% of the workforce, while the 40-49 and 50-59 age groups represented 26% and 27% respectively (OECD 2023a). Unfortunately, data subdivided by age and teaching subjects specifically for CS teachers were not readily accessible at the time of this report.

All the participants were teaching or studying computer science as a subject, with the following additional/second subjects: mathematics, physics, foreign languages, biology, chemistry, and other subjects. According to the pre-survey, mathematics and physics were the most common second subjects among the participants. Prior to the workshop, most participants had some experience teaching DL in class. As Figure 3 shows, while about half of the participants never taught AI in class, only five out of 23 respondents reported never conveying knowledge about DL in class. Yet, no respondents indicating teaching DL or AI as often as every week, with few (four respondents for DL and one for AI) reported teaching the content on the monthly basis. Out of 23

respondents, 13 reported teaching DL a few times a year and five participants reported doing so for AI.

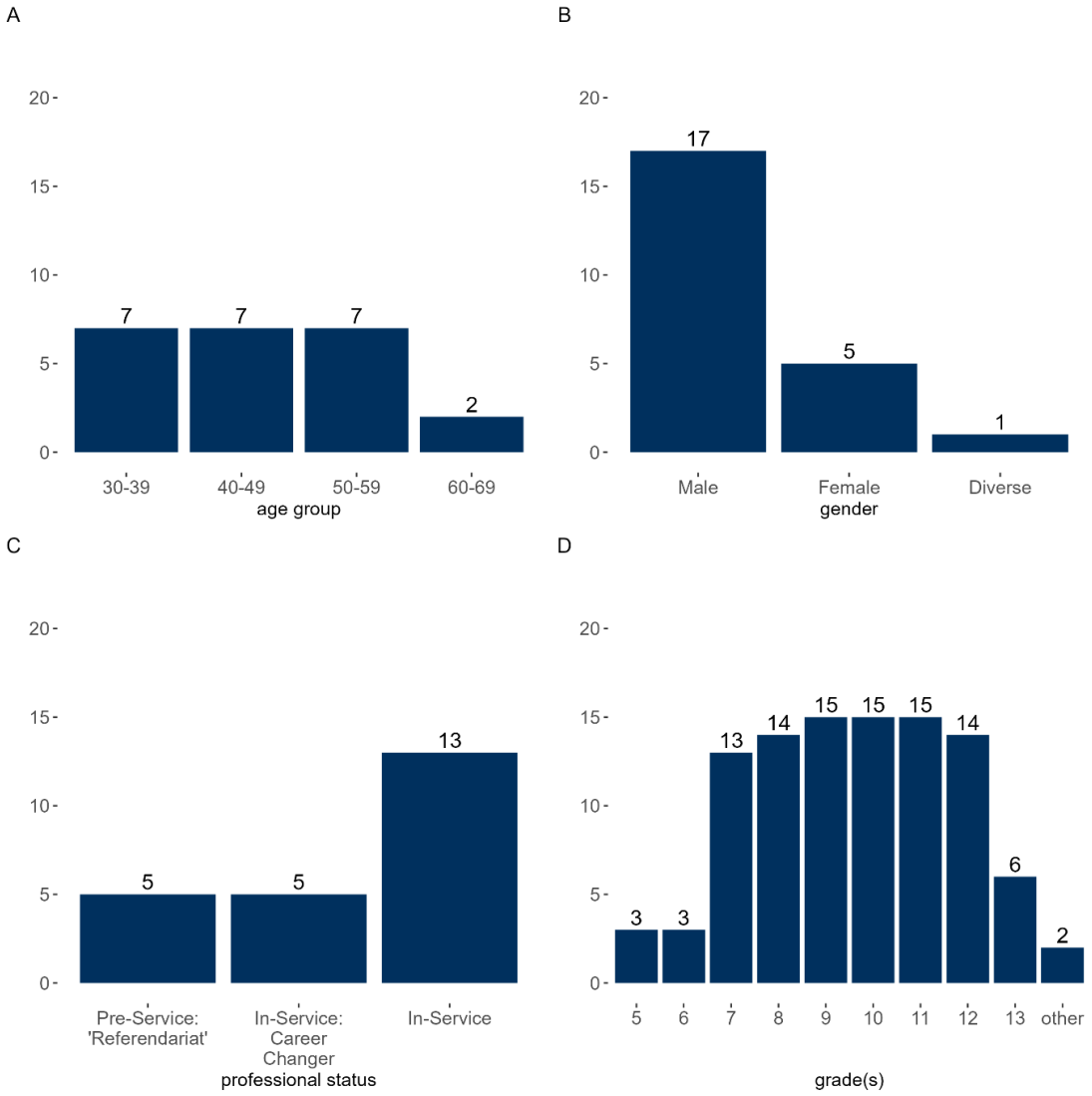


Figure 2 Socio-demographic data of the participants, 13.06.2022 Berlin, pre-evaluation survey, n=23

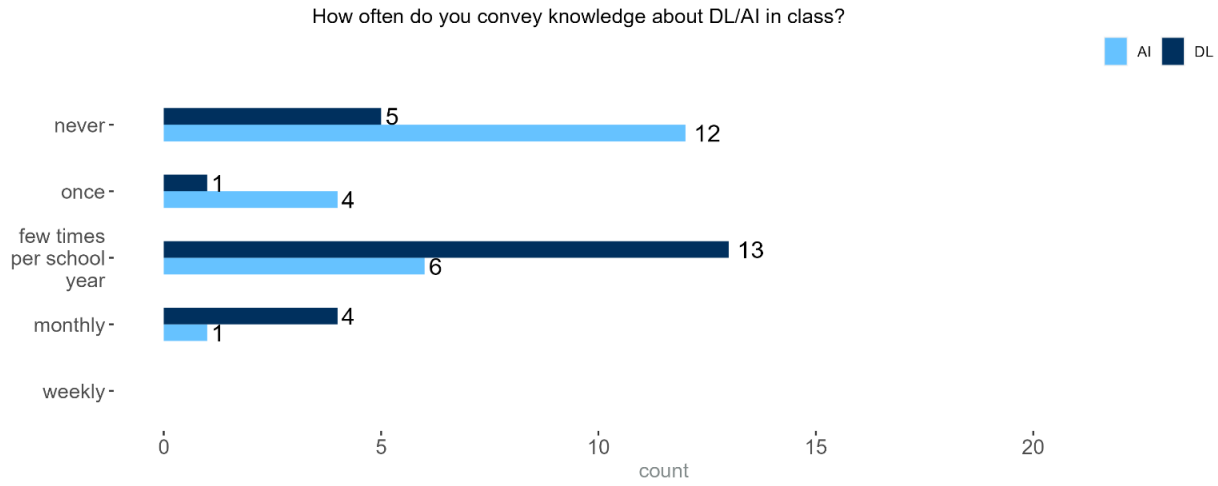


Figure 3 Experience with DL/AI, 13.06.2022 Berlin, pre-evaluation survey, n=23

2.2 Quantitative results

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

The survey item "I know how to use content about DL/AI in the classroom" measures perceived competences on how to use learned content in class on the 6-point scale. It was included in both pre- and post-evaluation surveys. As Figure 4 and Table 6 shows for both DL and AI items, on average respondents reported an increase in their perceived competences on how to use both topics in class.

For the DL part, participants' median score increased from an initial score of 2.0 to a post-training score of 4.0. The Wilcoxon signed-rank test indicated a statistically significant difference at the 1% level. Furthermore, the IQR decreased from 2.0 to 1.25, indicating that post-training values are less spread out and more consistent than the pre-training values. Similarly, for the AI part, participants' median score increased from an initial score of 2.0 to a post-training score of 5.0. The Wilcoxon signed-rank test also indicated that the two samples are significantly different at the 1% level. The IQR of 1.0 stayed the same, although the distribution of the data shifted upwards.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
I know how to use content about DL in the classroom	2.0	4.0	1.0	3.75	3.0	5.0	1	2	5	5
I know how to use content about AI in the classroom	2.0	5.0	2.0	4.0	3.0	5.0	1	3	6	6
I am willing to invest time and effort to incorporate DL into my teaching	-	5.0	-	3.75	-	5.25	-	2	-	6
I am willing to invest time and effort to incorporate AI into my teaching	-	5.5	-	4.75	-	6.0	-	2	-	6
After the training I have gathered enough competences to teach the learned content in class	-	4.0	-	3.0	-	4.25	-	2		5

Table 6 Summary statistics of pre- and post-results for the survey items on teachers' perceived competences to use DL and AI content in class and post-survey results for additional items, 13.06.2022 Berlin, pre- and post-evaluation survey, n=20

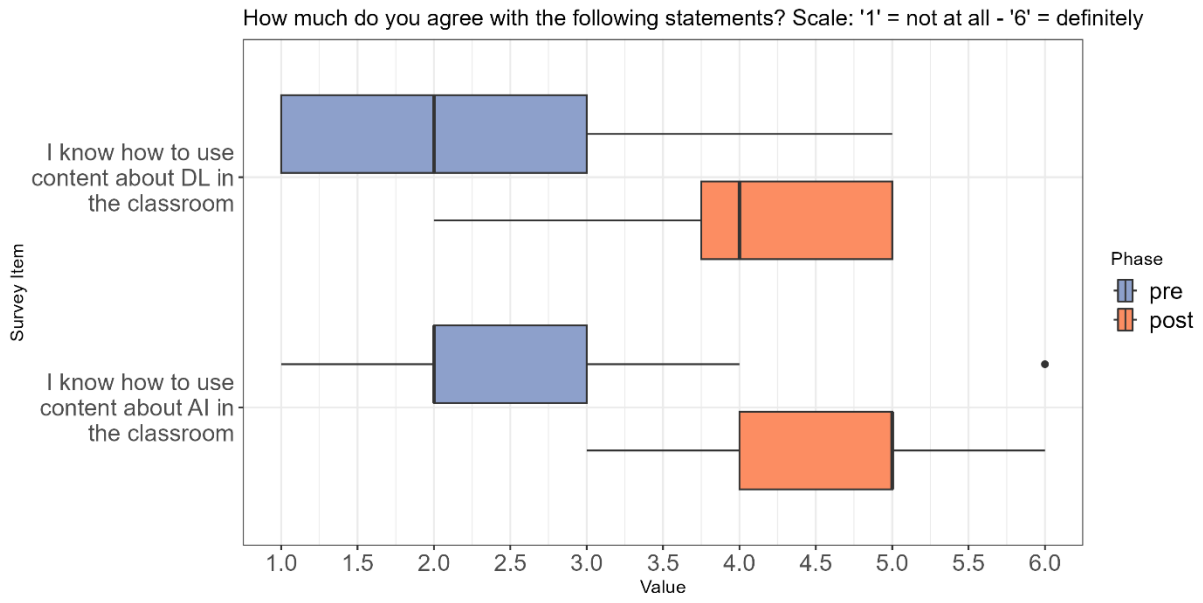


Figure 4 Boxplot comparison of pre- and post-results for the survey items on teachers' perceived competences to use DL and AI content in class, 13.06.2022 Berlin, pre- and post-evaluation survey, n=20

Figure 5 and Table 6 present additional survey items that were exclusively measured in the post-survey. Participants were asked to indicate their level of agreement with the statements: "I am willing to invest time and effort to incorporate DL/AI into

my teaching." For both DL and AI items, most participants selected categories "5" and "6," with "6" representing "strongly agree." The results indicate that, on average, participants expressed a strong willingness to invest time and effort to incorporate both DL and AI into their teaching, with a higher willingness observed for AI compared to DL.

Furthermore, participants were asked to what extent they agree with the statement: "After the training, I have gathered enough competences to teach the learned content in class." Overall, the statistics for this post-survey item indicate a moderate level of agreement among participants, with a concentration of responses around the value of 4.0 with the middle 50% of the responses ranging between 3.0 and 4.25.

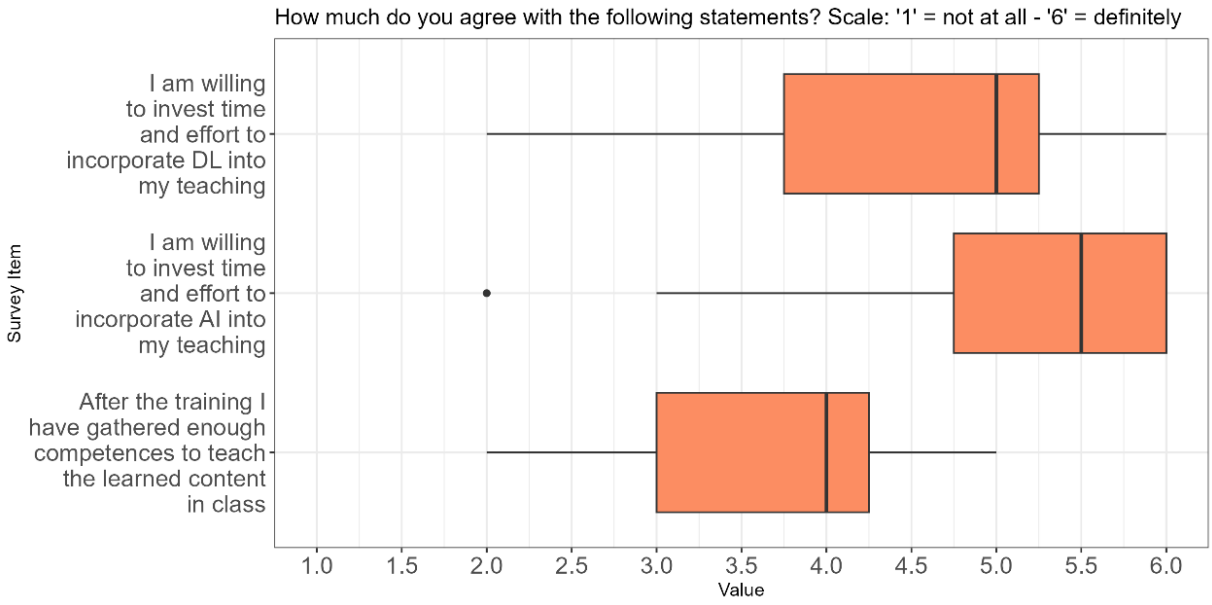


Figure 5 Boxplot of post results for the survey items on teachers' perceived competences and willingness to invest time and effort to incorporate DL and AI into their teaching, 13.06.2022 Berlin, post-evaluation survey, n=20

Overall, the findings indicate that the CS training in Berlin, Germany on average enhanced teachers' perceived level of competences on how to use DL and AI in their teaching. In addition, the post-survey measures reveal a positive inclination among participants to dedicate time and effort towards integrating DL and AI into their teaching. However, there was only moderate consensus among participants regarding

their perceived competences to teach the learned content in class, with a notable concentration of responses around a moderate rating of 4.0.

2.2.2 Understanding of DL/AI concepts introduced in the training

To measure gain in understanding of DL and AI concepts introduced in the training, we have used both self-report measures and objective knowledge questions in the pre- and post-self-assessment and knowledge tests administered separately from the surveys.

For the self-report knowledge questions (rated on a scale from 1 - "strongly disagree" to 5 - "strongly agree"), there is some improvement in the post-training median scores, with a larger increase observed for the AI set of questions (see Figure 6 and Table 7). The median score for DL increased from 2.0 to 3.0, while the scores for AI increased from 2.0 to 4.0. The Wilcoxon signed-rank test indicates a statistically significant difference between the pre-and post-samples for the DL and AI questions at 5 and 1% level, respectively. The spread of the middle 50% of the data reduced for both DL and AI questions, albeit more for the AI questions. Considering the self-assessment scale, which ranges from 1 (strongly disagree) to 5 (strongly agree), and the median post-results of 3.0 for DL and 4.0 for AI, the findings suggest that there is potential for improvement, especially for the dimension of DL which primarily included the topic of data lifecycle.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Self-assessment for DL	2.0	3.0	1.75	3.0	3.0	4.0	1	2	4	4
Self-assessment for AI	2.0	4.0	2.0	3.88	3.0	4.0	1	3	4	5
Knowledge test DL	2.0	2.5	1.65	1.0	2.4	3.13	0.0	0.4	2.8	4.0
Knowledge test AI	4.2	6.0	3.25	4.88	6.25	7.53	1.0	3.5	8.5	9,1

Table 7 Summary statistics of pre- and post-results for the self-assessment and knowledge test, 13.06.2022 Berlin, pre- and post-self-assessment and knowledge test, n=16

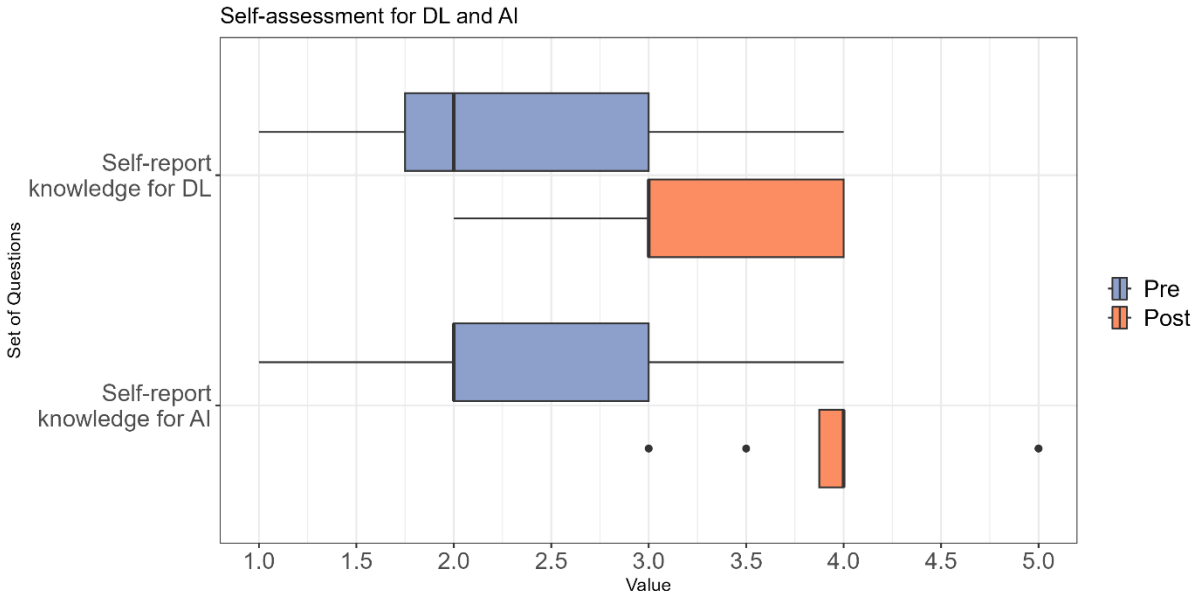


Figure 6 Boxplot comparison of pre- and post-results for self-report knowledge questions on DL (15 questions) and AI (10 questions), 13.06.2022 Berlin, pre- and post-self-assessment and knowledge test, n=16

Objective knowledge tests were conducted to assess participants' knowledge in both DL and AI dimensions (see Figure 7 and Table 7). Immediately after the workshop, participants demonstrated a higher median of 6.0 for the AI part compared to the pre-workshop, which constituted 4.2. The Wilcoxon signed-rank test indicated a statistically significant difference between the pre-and post-samples at the 1% level, indicating an improvement in participants' knowledge of AI concepts. Although the median score increased, the IQR did not change. Also given that the maximum score was 10.0, we cannot conclude that on average participants managed to grasp all the concepts introduced in the workshop.

There were no statistically significant differences in the pre- and post-results for the DL part. This implies that on average participants' knowledge of DL concepts remained relatively unchanged in this sample when comparing before and after the workshop. The descriptive statistics suggest that there is greater variability in participants' understanding of DL concepts after the training, with some participants showing a higher level of knowledge and others demonstrating a lower level. IQR increased

from 0.75 to 2.13, indicating that the spread of the central 50% of the data became larger.

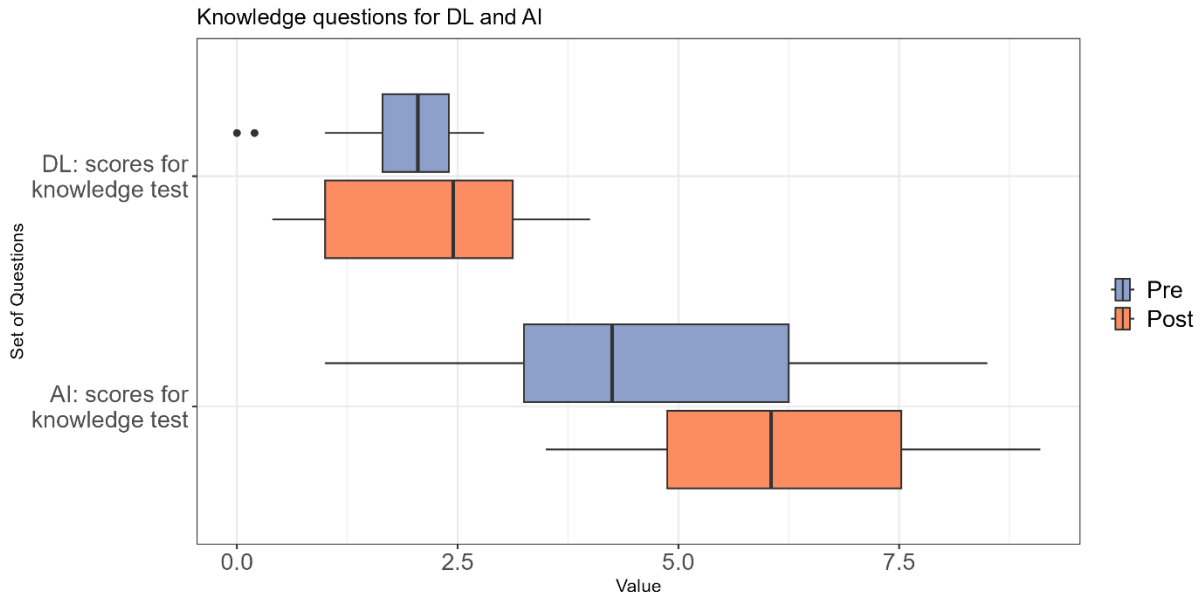


Figure 7 Boxplot comparison of pre- and post-results for knowledge questions on DL (4 questions) and AI (10 questions), 13.06.2022 Berlin, pre- and post-self-assessment and knowledge test, n=16

2.2.3 Attitudes towards DL/AI

Table 8 and Figure 8 present the distribution of survey items related to teachers' attitudes towards DL and AI with a focus on their perception of the importance of both topics. These items were measured in the pre- and post-policy experimentation surveys: "In the future, teaching DL/AI will provide added value to students" and "I think the content of DL/AI is missing in the current framework curriculum."

Prior to the training, participants held varying opinions regarding the added value of teaching DL and AI to students, with a median score of 4.0 for DL and 4.5 for AI and a quite large IQR. While the IQR notably decreased for the post-scores of the DL item, the median remained unchanged. In contrast, for the post-item on AI, participants reported a higher score, with a median of 6.0. The Wilcoxon signed-rank test indicated a statistically significant difference between the pre- and post-training AI values at the 5% level.

Regarding the perception of DL/AI content in the current framework curriculum for CS, participants leaned slightly towards the view that it is missing, with a consistent median score of 4.0. There is no statistically significant difference between the pre- and post-results. Both the pre- and post-results showed a relatively wide spread of the data. The median for the AI curriculum item slightly decreased after the training from 4.5 to 4.0.

An additional post-survey item on the societal importance of DL and AI indicated a high level of agreement with the statement that DL and AI have enough societal importance to be incorporated into the curriculum (see Figure 9 and Table 8).

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
In the future, teaching DL will provide added value to students	4.0	4.0	3.0	4.0	6.0	5.0	2	3	6	6
In the future, teaching AI will provide added value to students	4.5	6.0	3.25	5.0	6.0	6.0	1	3	6	6
I think the content of DL is missing in the current framework curriculum of computer science	4.0	4.0	3.25	3.0	6.0	5.0	2	1	6	6
I think the content of AI is missing in the current framework curriculum of computer science	4.5	4.0	4.0	3.0	5.75	5.0	2	1	6	6
The topic of DL is of enough societal importance to integrate it into the curriculum	-	5.2	-	5.0	-	5.0	-	3	-	6
The topic of AI is of enough societal importance to integrate it into the curriculum	-	5.5	-	6.0	-	6.0	-	3	-	6

Table 8 Summary statistics of pre- and post-results for a set of survey items on teachers' perception of importance of DL and AI, 13.06.2022 Berlin, pre- and post-evaluation survey, n=18-21

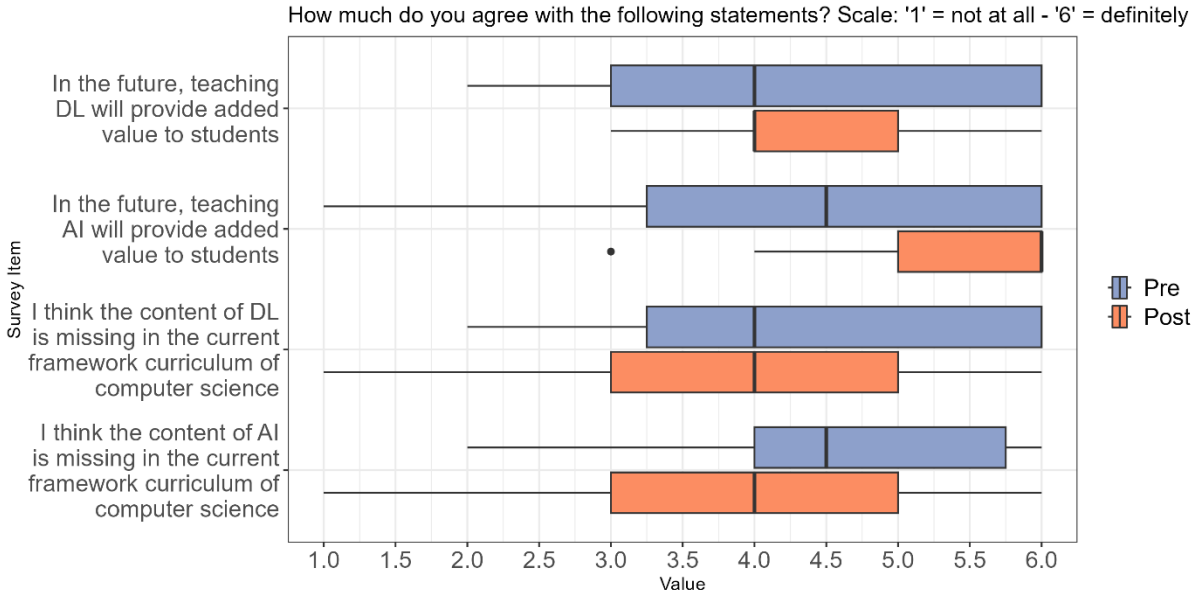


Figure 8 Boxplot comparison of pre and post results for survey items on teachers' perception of importance of DL and AI, 13.06.2022 Berlin, pre- and post-evaluation surveys, n=18

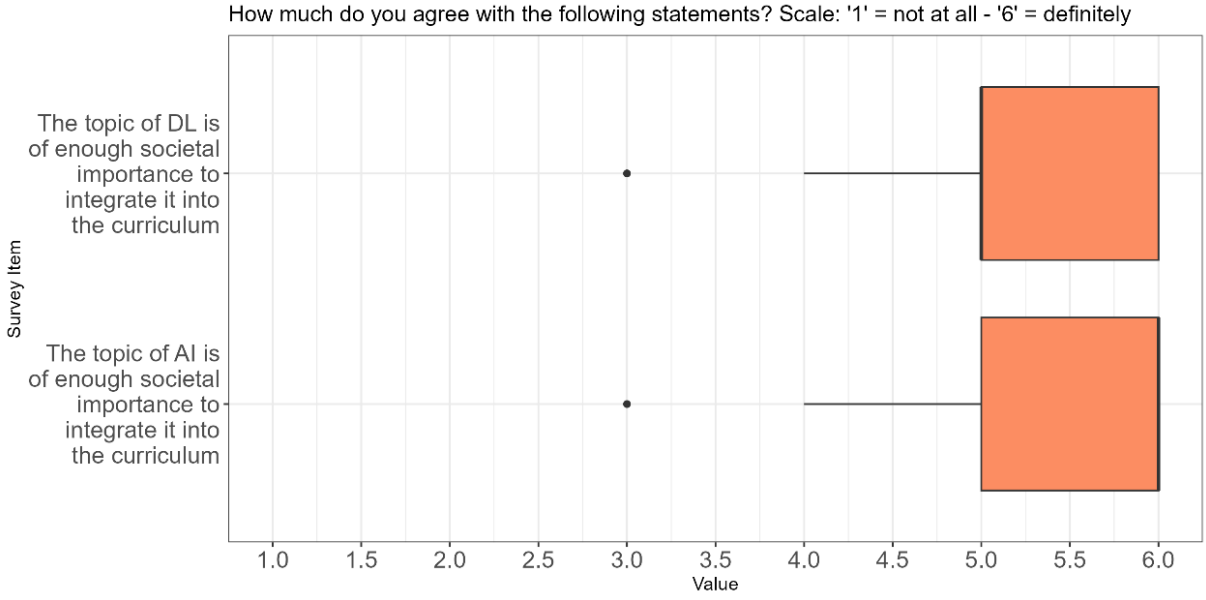


Figure 9 Boxplot comparison of post results for the survey item on teachers' perception of societal importance of DL and AI for integration into the curriculum, 13.06.2022 Berlin, post-evaluation survey, n=21

2.2.4 Perceptions of student engagement

To explore teachers' expectations regarding students' interest and their perceived ability to generate student interest and engagement in the topics of DL and AI, the post-survey included the following survey items: "I can imagine that my students will show enthusiasm for the overall topic of DL/AI" and "I am confident that I can get students excited about DL/AI-projects".

As depicted in Figure 10 and Table 9, the findings reveal generally positive expectations regarding students' interest, particularly in relation to AI. However, it is worth noting that teachers rated the expected student enthusiasm for DL topics lower compared to the same item for AI topics. This difference suggests that on average teachers anticipate greater student engagement and excitement when it comes to AI compared to DL.

Figure 11 and Table 9 also indicate a notably high level of confidence among teachers regarding their ability to motivate students for AI projects. This suggests that teachers generally feel well-equipped and self-assured when it comes to fostering student engagement and motivation in AI-related projects. In contrast, teachers' perception regarding DL projects falls within a more moderate range.

Question	Median	Q1	Q3	Min.	Max.	Scale
I can imagine my students show enthusiasm for the overall topic of DL	4.0	3.0	5.0	2	6	1-6
I can imagine my students show enthusiasm for the overall topic of AI	6.0	5.0	6.0	3	6	1-6
I am confident that I can get students excited about DL-projects	3.0	3.0	3.0	2	4	1-4
I am confident that I can get students excited about AI-projects	4.0	3.0	4.0	2	4	1-4

Table 9 Summary statistics of post-survey results for items on perception of student engagement, 13.06.2022 Berlin, post-evaluation survey, n=21

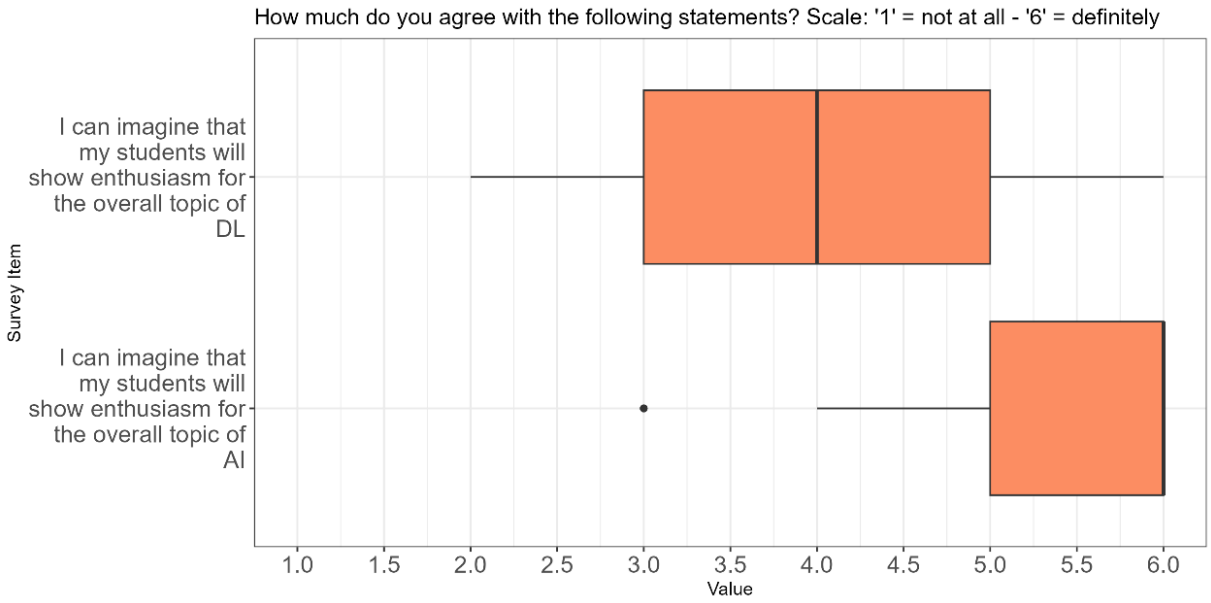


Figure 10 Boxplot of post results for the survey item "I can imagine that my students will show enthusiasm for the overall topic of DL/AI", 13.06.2022 Berlin, post-evaluation survey, n=21

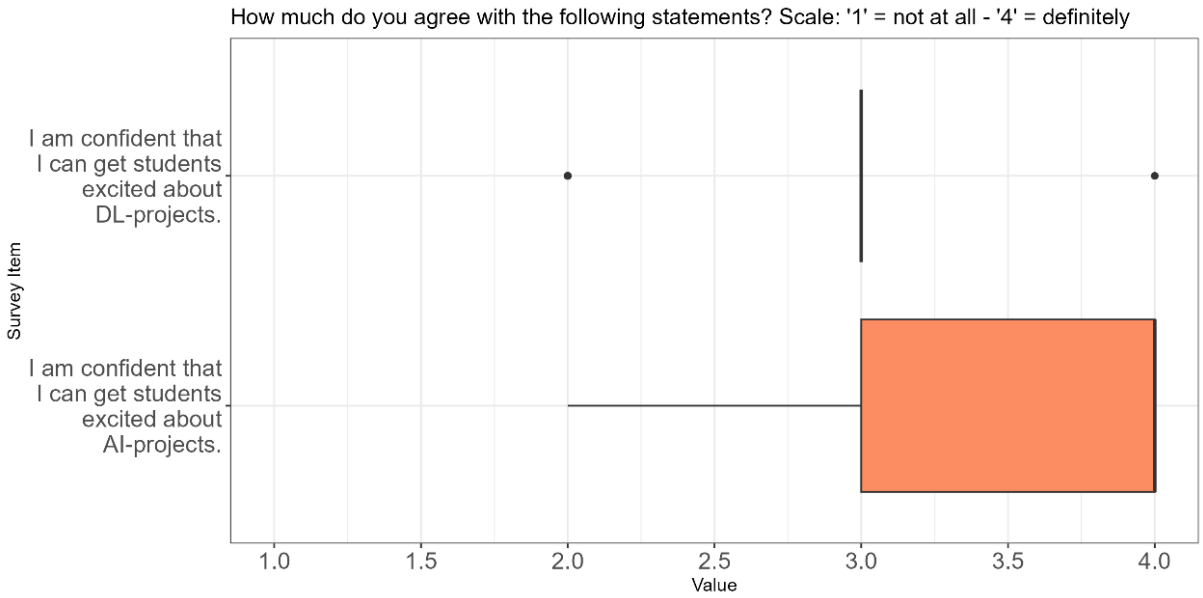


Figure 11 Boxplot of post results for the survey item "I am confident that I can get students excited about DL/AI-projects", 13.06.2022 Berlin, post-evaluation survey, n=21

2.2.5 Feedback on the learned content and format

The post-evaluation survey included a series of questions to assess participants' reactions to the topics and materials/exercises covered in the training. Figure 12 and

Figure 13 present the results of this assessment. The findings indicate that most participants consider the presented topics of classical AI and machine learning (ML) to be highly suitable for their teaching. Approximately 80% of respondents selected a score of "5" or "6" (where "6" indicates a high level of suitability). However, the topic of data lifecycle was viewed as less suitable, with only 50% of respondents selecting a score of "5" or "6." Additionally, 15% of participants chose a score of "2" or "1" where "1" indicates that the topic is not suitable at all. Similarly, when evaluating the materials and exercises used in the training, it was observed that the exercise on data lifecycle with Orange3 tool did not yield as positive results as the other exercises. While at least 70% of respondents rated all other exercises as "5" or "6" indicating high suitability for teaching, only 20% of participants gave the data literacy exercise with Orange3 such a high rating.

These findings suggest that the topics of classical AI and ML were well-received by participants in this sample. However, there is room for improvement in the delivery and effectiveness of the data lifecycle topic, particularly the use of Orange3.

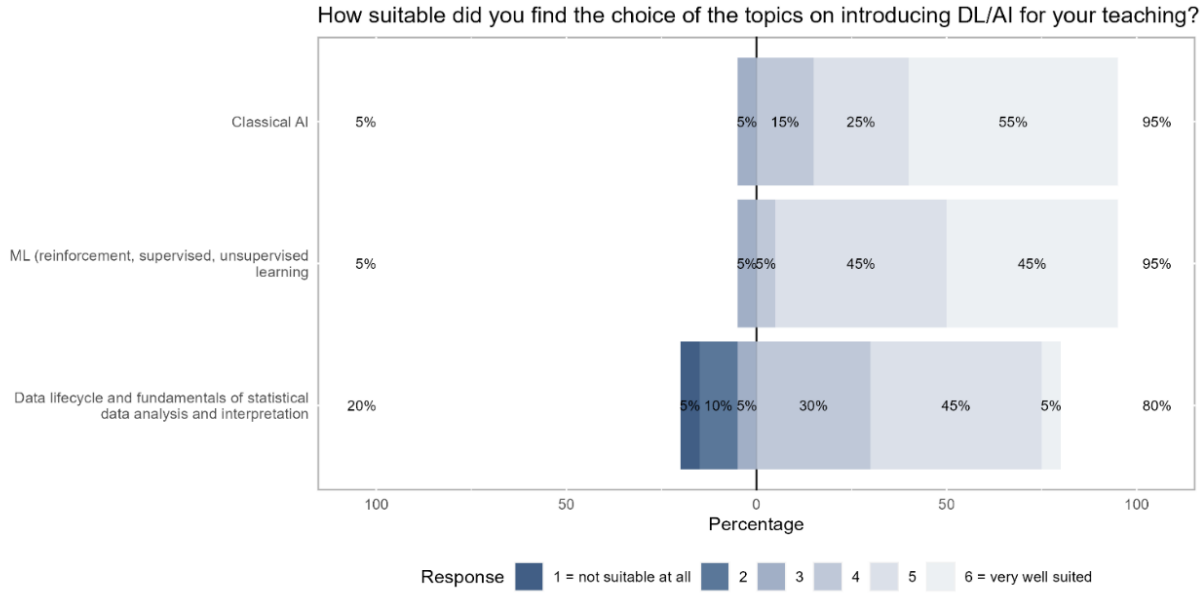


Figure 12 Distribution of the post results for the assessment of the training topics, 13.06.2022 Berlin, post-evaluation survey, n=20

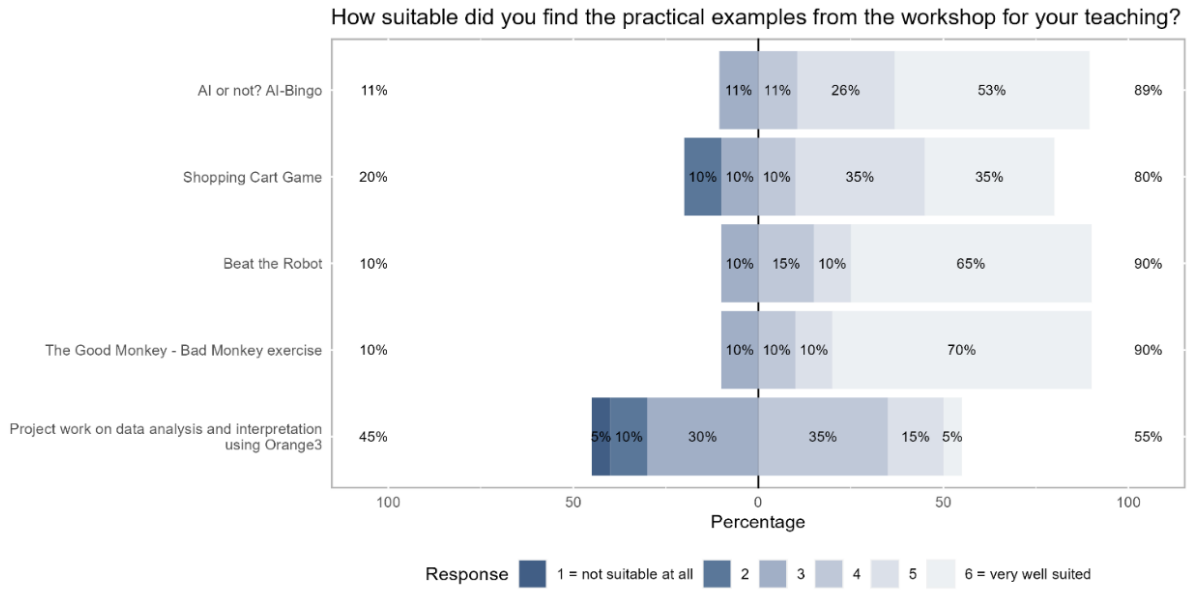


Figure 13 Distribution of the post results for the assessment of the training materials/exercises, 13.06.2022 Berlin, post-evaluation survey, n=20

Figure 14 shows additional items on teachers' feedback, including training format and length. With respect to the length of the training, participants are somewhat divided in their opinions, and there is no indication of strong consensus for the statement that the workshop should have lasted longer. Of the 20 participants, eight selected the higher scores of "5" and "6", whereas six respondents opted for the lower scores of "1" and "2". The remaining participants selected scores within the intermediate range. The results regarding the item "I wish more content would be covered in the training" are also rather mixed, with 50% of the respondents strongly leaning towards disagreeing with this statement and 30% of the respondents strongly agreeing with it. There is, however, a strong consensus that the training did not have too many teaching materials and that the interactive format was appropriate. Most of the participants lean towards agreeing with the statement, "The training showed me which competences I lack to teach the relevant content in class". The statement "I would need a lot of preparation to teach the topics and application examples in class" had a rather mixed response, slightly leaning towards agreeing with the statement.

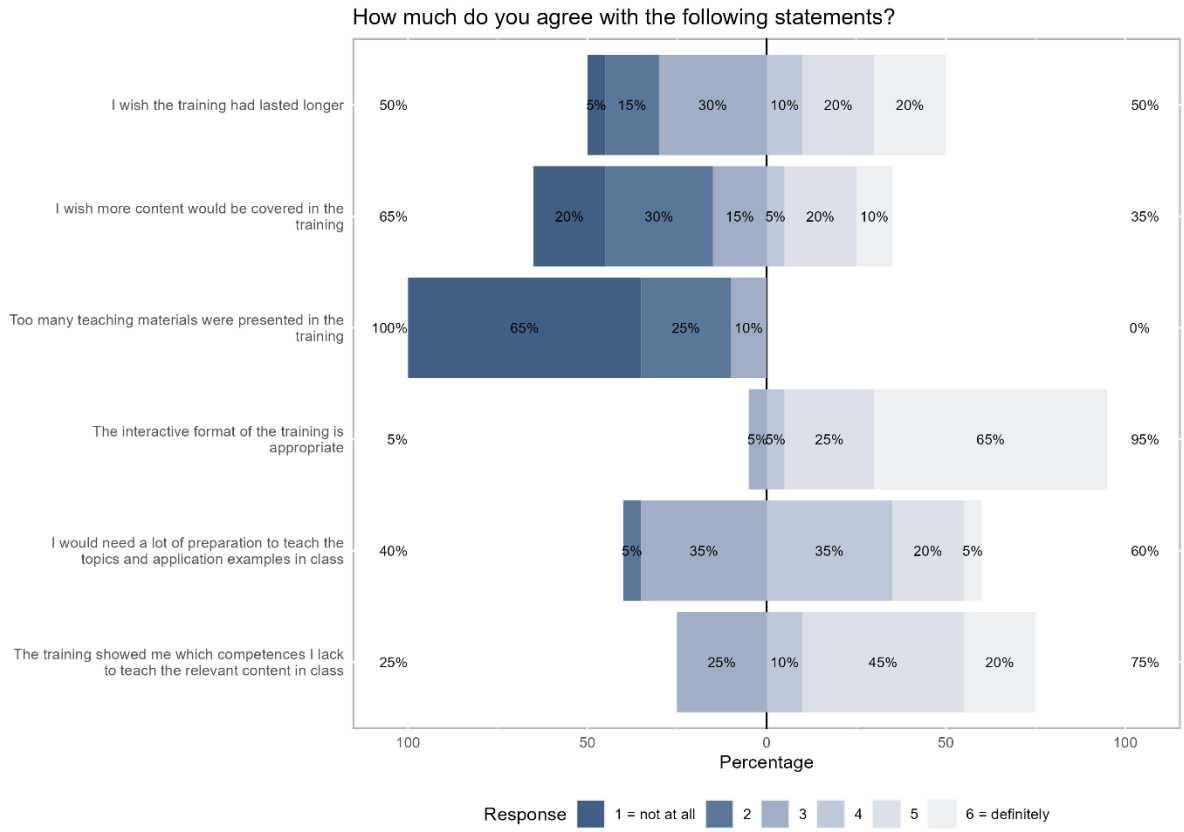


Figure 14 Distribution of the post results for the assessment of the training format and outcomes, 13.06.2022 Berlin, post-evaluation survey, n=20

2.2.6 Follow-up

Approximately six months after the intervention, we asked participants to complete a follow-up survey, primarily to ascertain if they had managed to integrate the content learned during the training into their classes. The survey was completed by seven participants. Out of these participants, three reported teaching the content on DL and AI in their classes, albeit two of them had already taught DL and AI prior to the training. All four individuals who indicated that they had not yet begun teaching DL or AI content at the time of the follow-up survey reported plans to do so. Among the topics that they plan to use in the future are: AI in everyday life (e.g., "AI Bingo"); machine learning (e.g., robot chess game "Beat the robot", "Good Monkey - Bad Monkey", and "Shopping cart" game); classical AI (e.g., "Beat the robot"); image generators and

prompts; deep learning; and programming with Python and TensorFlow. Nobody indicated plans to teach the topic "data lifecycle and the basics of statistical data analysis and interpretation" (e.g., Orange3).

2.3 Qualitative results

2.3.1 Personal interviews

The themes listed below were deductively derived from the research questions and the structure of the interview guide. The annotations on the right refer to categories that were assigned inductively throughout our qualitative analysis process. Given the novel nature of the research questions, we chose not to limit our analysis to a purely deductive approach. Though our initial research questions and the interview guide shaped the general themes, it was the inductive analysis of participant responses that filled these themes with detailed and meaningful content.

A. Training:

Factors prompting training participation (prior to the training)

Information about the event was obtained via suggestions from superiors, colleagues, as well as the faculty of didactics of Computer Science at the Freie Universität Berlin. Further resources, such as participation in other continuing education courses and the mailing list from the GI Fachgruppe IBBB, also proved useful in obtaining details about potential training opportunities.

Information pertaining to the training was obtained from external sources

Regarding individual initiative, it should be noted that some participants undertook active online searches for the training, frequently utilizing "computer science" as a key search term.

One person actively searched for the training online

Expectations (prior to the training)

In the realm of DL and AI, the expressed expectation of participants was to gain a foundational understanding. Regarding Orange3, the aim was to bolster professional competences. In the context of AI, the emphasis extended beyond theoretical knowledge to include tangible examples, thereby enabling a more effective contextualization of familiar content.

Obtain foundational understanding and get introduced to practical examples

For both DL and AI classes, the expectation was to apply the in-service training materials and to receive both content-related and didactic input. In addition, the teachers wanted to receive teaching materials to support them. In the specific area of DL, there is a wish to get access to this topic and to get ideas about how to design classroom lessons with Orange3. With respect to the AI teachers would like to have support for designing individual lessons, including the content, sequence and the structure: *"Especially didactically in which order is it best to do it and what exactly belongs where in such a larger process of a teaching series."* (Participant B - 13.06.22, Berlin)

Teachers wished for (ready-to-use) teaching materials and didactic input on how to structure and sequence lessons

Design according to familiarity with DL and AI

Both for moderate/average levels of familiarity in DL, Orange3, and AI, as well as for no familiarity in Orange3, the design of the training was tailored well to the levels of familiarity of the teacher.

Training well-aligned with pre-existing knowledge

Difficulty of the topics for the teachers themselves

The training provided basic knowledge and understanding of the topics of DL and AI. However, due to the sheer breadth and depth of these subjects, mastering them fully poses a significant challenge. Acquiring them on the job is also difficult, as numerous training courses would be required. Although some basic knowledge could

Basic knowledge was acquired, but mastery of DL and AI remains challenging

be acquired after further training, there is still a clear difference to expertise in this field.

In the area of artificial intelligence (AI), the use of examples/exercises contributed to better understanding: *"This introduction and the various forms of learning, that they were explained again, the three [types of machine learning], that was very important for me. And there were a few difficulties in really understanding the text. That became very clear to me with the examples."* (Participant D - 13.06.22, Berlin) Regarding AI, the high complexity of the topic became clear, especially in relation to machine learning and neural networks. Understanding at a deeper level is assessed as more challenging than the application itself. In particular, the functioning of image recognition (software) remains a black box for teachers and students, i.e., not fully comprehensible: *"Sure, if I have lots of pictures of cats, then I might recognise a cat. But how that works exactly often remains a black box at that point, both for the students and for me."* (Participant F - 13.06.22, Berlin)

The complexity of AI, particularly machine learning and neural networks, became evident. Image recognition Software's inner workings remain a "black box"

The area of data analysis with Orange3 proved to be exceptionally complex. Notably, the process of creating the decision tree remained obscure, often perceived as a "black box". Therefore, proficient use of Orange3 demands additional personal engagement and commitment.

Orange3 and decision tree are seen as complex - he black box perception

B. Teaching DL and AI:

Integration of DL and AI into teaching CS at the secondary level (prior to the training)

In the field of computer science, DL was explored primarily through work with databases. This work encompassed various aspects of data management, including data collection, definition, cleansing, archiving, preparation, and visualisation of data sources.

Different topics of DL (via databases) and AI (via robotics) were integrated

The topics of AI, especially in the area of robotics, was dealt with either once or continuously in a series of lessons.

Experience with integration of DL and AI into teaching CS at the secondary level (prior to the training)

Successful integration of AI via robotics, but as topics became more complex, integration proved more challenging

The following experiences in the field of AI were collected in the tenth-grade computer science class. The use of unplugged materials in class has proved to be successful. This approach has motivated the students and encouraged their sense of discovery. It is a form of exploratory learning where, subsequently, there is the opportunity to discuss the topics covered. However, it was found that with increasing complexity and depth of the topics, such as practical and theoretical work with neural networks, the degree of difficulty increases. Therefore, more extensive projects could not be carried out. In addition, the assessment of student performance is difficult. This concerns both the problem of pure memorisation on the part of the students and the actual conception of the performance assessment: *"[...] I then had them write an exam on the topic, and I found that a great challenge, to see if I now explained to them on a handout where neural networks are explained, how can an exam on this be constructed"* (Participant B - 13.06.22, Berlin)

Frustration or disappointment when students confronted the complexity

Overall, positive experiences were gained during the implementation of DL and AI. There was a noticeable interest among the students. However, a sense of disillusionment has been recognised: *"Students are very interested in the subject at first, but then they are also really disappointed sometimes when they can't get into it that much because they don't grasp the complexity."* (Participant F - 13.06.22, Berlin)

There was no proactive inquiry from the students prior to teaching topics of AI. The integration of AI into the classroom was feasible without the need to replace other topics.

Teaching AI was possible without replacing other topics

The teaching of DL was well received by the students.

DL was well received

Ability of (more) effective integration (after the training)

The starting point for lesson planning in DL and AI is rated as good.

Training as a good entry point

One teacher reported his plan to use Orange3 for the topic of AI rather than treating it solely as a statistical subject. To ensure effective integration, coordinating discussions with colleagues about the teaching content are necessary. However, a direct implementation of the training content and exercises may not be immediately possible: "What I have experienced today, I will not be able to go one-to-one into the classroom and do that without further engagement [with the training's content], but approximately seventy to eighty percent of the way I have gone today." (Participant B - 13.06.22, Berlin)

Orange3 can be integrated as an AI supporting tool

The integration of AI into teaching is a key objective, with the content from the in-service training offering suitable material for adaptation. The first contents, especially classical AI, will be integrated into the robotics lessons immediately. In the upcoming lessons on the topic of smart home, AI can be integrated further. There is a positive expectation regarding the integration of the topic with the aim of long-term inclusion in computer science education. However, a prerequisite for this is the creation of an implementation plan to achieve this goal. It is planned to conduct a short-week teaching unit on AI for the students. A longer series of lessons (spanning 4-6

AI will be added to robotics lessons immediately. A short teaching unit dedicated to AI is envisioned, but a full 4-6 weeks of only AI might be challenging

weeks) solely dedicated to AI can be challenging. Therefore, integrating these concepts into related subjects, like smart home technology, proves to be a more feasible strategy.

Prior materials for introducing AI were heavy on theory, lacking motivation for students

In terms of material, previous attempts to introduce AI have not been easy to implement: *"It was very theoretically loaded. You have a video, media that you can look at, what you need it for, but exactly this approach, that was also still missing to motivate the students."* (Participant C - 13.06.22, Berlin)

"Good Monkey - Bad Monkey" game and "AI Bingo": 1:1 implementation is not possible

The exercise "Good Monkey - Bad Monkey" is mentioned as a suitable approach. It is planned to design own material and to use both modified (e.g., with Lego bricks) and unmodified approaches (for tenth grade). However, a 1:1 implementation is not possible without further self-study and refining existing materials, for example when using the exercise "AI Bingo".

Teachers gained insights on how to introduce AI, the structure to follow in lessons, and received a boost in motivation. Thematic decisions, such as what specific AI topics to cover and at which grade level, still need to be made

When integrating DL and AI, simple questions were clarified in the training and preparation is now possible. After the workshop, ideas for introduction, approaches (e.g., through exercises) and the basic structure in class were gained, which increases the teacher's motivation as well as enables a better didactic and more motivating implementation. Before integration, however, some points need to be determined, such as the thematic focus and the grade level in which the implementation takes place. For a teaching series of 6 to 8 weeks, further engagement with the training content is necessary, including for the upper secondary level (grade 11-13), where the topics should be explored in greater depth. Therefore, one is not bound to the training material: *"And I can now optimise the material I developed for*

myself, based on material that was already there, once again with the material I took with me today." (Participant F - 13.06.22, Berlin)

Integration of AI without DL is deemed challenging

It is further stated that a separate integration of AI (without DL) is considered inconceivable. It was perceived as too detached from DL in the training and comes along with a higher level of difficulty.

AI can be introduced even in a single lesson focusing on simple methods

It is easily imaginable to have a single lesson dedicated to teaching introductory simple methods about AI. The integration is expected to take place in the upcoming school year. To effectively convey the subject matter, it may be beneficial to consult textbooks from the respective federal state.

DL can be taught through teaching databases in a concise manner

Integration of the DL subject via teaching databases is possible in a compressed teaching block. After the training, there will be an evaluation and reflection of Orange3.

Exercises from the training, like the "Beat the robot" game, were seen as good introductory points, but sometimes require modifications, such as using felt-tip pens instead of stones

The exercises from the TrainDL training were evaluated as a very good thematic introductory/entry point and could be implemented with well-prepared material. Previous attempts to implement AI in other areas outside of robotics often failed due to a lack of material or approach. However, these difficulties were successfully overcome through exercises during the training. The modification of the material (e.g., using the "Beat the robot" game as an example) is not always directly applicable in a 1:1 manner, but it could be adapted, for instance, by using felt-tip pens instead of stones.

Time is needed for material creation and preparation

The creation of tasks and material takes time, but it could be achieved by using the training material. However, the processing of

the material for exercises also requires time, e.g., laminating the documents.

Difficulty of conveying knowledge to school students (after the training)

In computer science, access to AI for students is possible through exercises, materials, and teacher support.

Integrating DL and AI content into the school curriculum is plausible and can be customized based on students' grade levels

Regarding DL and AI, the content, examples, and materials provided can be integrated in in the framework curriculum. Integration may vary depending on the grade level, with older grades that can handle more challenging examples. Even at lower grade levels with a good degree of proficiency, the content could be taught. Overall, it would remain a challenge, but would be feasible.

DL/Orange3 suitable for higher grade levels

DL is more suitable for an advanced course rather than a basic course (upper school), as it involves more scientific work. Orange3 as well, especially if they already have experience with similar programs. However, it should be covered in more detail than in the TrainDL training and taught live to the students: *"Or click together with them simultaneously because I think that there will be quite a lot of questions. We had already such an experience [during the exercise in the TrainDL training] - Why don't I see anything. [...] You have to add this and that first. So that's probably not something for students to work with by themselves or you really have to walk around constantly, to help them."* (Participant E - 13.06.22, Berlin)

AI exercises as a good introduction and can be employed for a wide range of students

There are also positive aspects around AI: The exercises from the training with suitable materials are also accessible for relatively weak students. Especially the initial exercises and examples are optimal. The "Beat the robot" game is well suited for intermediate and

advanced levels up to grade level 13. However, challenges are also seen, but these are mitigated by the examples. For some classes, the use of unplugged examples with loose materials (stones, smarties) would not be advisable.

C. Establishment and steps to integrate DL and AI:

Framework curriculum integration

The linking of DL and AI is seen as a relevant topic area in the framework curriculum. This is because these topics have societal as well as political implications: *"[...] the societal implications directly affect almost every student, as they receive suggestions for various things on platforms like TikTok or YouTube."* (Participant A - 13.06.22, Berlin)

DL and AI are seen as relevant topics

The students must be able to reflect critically on such content. The framework curriculum ensures contact with knowledge about the processes in a technologised society.

Students must develop the ability to critically assess such content

The inclusion of DL is not considered necessary, as not all students need to be familiar with the topic. It is suggested that it should be integrated in tertiary education instead of including it in secondary education. AI is considered more relevant for inclusion in the framework curriculum than DL, since AI is more student-oriented and less abstract than DL, which in turn is more scientifically pronounced than AI. In addition, databases (DL) are already integrated in the framework curriculum.

AI has higher priority for secondary level than DL

On the other hand, the inclusion of AI is necessary, but it is also suggested to integrate in other related subjects. The inclusion of the topic enables students to develop the ability to participate in opinion-forming and democracy-building. AI has an increasing relevance

With the introduction of AI, it's vital to outline specific competencies. This might necessitate adjustments in other

and people encounter it unconsciously in everyday life. It is important that students understand this topic. When including AI, concrete topics, as well as competences, should be defined. Subject-specific and cross-curricular considerations must be made to possibly eliminate other content: *"Because the instruction time will not become longer, the question arises: What should be removed elsewhere? Now, for example, in the semester with the specialisation areas, there is theoretical computer science and perhaps we should discuss whether it should remain as extensive as it is, or whether we can't say that parts of what is learned there can also be learned when working on modelling for AI."* (Participant B - 13.06.22, Berlin)

areas of the curriculum to accommodate AI, prompting to consider what might be removed or reduced in terms of content

It is emphasised that it is important to include AI in Berlin's curriculum as soon as possible [Note: all training participants work in Berlin], as other federal states such as Baden-Württemberg have already integrated AI in grade levels five to six.

It is important to integrate AI in Berlin's framework curriculum

Regarding the placement of the topics, DL can be taught not only in computer science but also in other subjects, e.g., mathematics (specifically statistics), natural sciences, or social sciences.

DL can be integrated not only in CS, but in other subjects

Furthermore, data structuring and cleaning are considered unimportant, while chart interpretation, reading statistics, evaluation, data preparation and database systems are considered important.

DL should focus on evaluation and interpretation of data, rather than cleaning and structuring

AI goes beyond computer science teaching and is preferably regarded as a compulsory course of CS teaching or should at least be integrated via the elective subject area.

AI can be introduced as a core component of computer science education or be woven into elective subjects.

In the context of AI, it is emphasised that students should be able to name examples and risks. Other topics that should be covered include societal aspects, spam filters, image recognition as well as other *"old classics"* (Participant C - 13.06.22, Berlin) in the field of AI.

Students should be familiar with practical examples of AI applications and understand the associated risks.

Steps for anchoring in the classroom

The current CS framework curriculum for Berlin is imprecise and difficult to follow. A new one is a prerequisite for anchoring the topics and has a steering function: *"It should first be included in the framework curriculum, with even more subtopics that clearly explain what is meant by it, i. e. societal aspects, what is programmed in relation to it, and which algorithms are to be addressed."* (Participant D - 13.06.22, Berlin)

Revised framework curriculum is needed

The repetition cycle must then also be clarified: e.g., a recurring spiral curriculum or a separate module in a specific grade level. An early treatment of these topics is beneficial: *"Because actually, it would make sense to do it at least once in secondary level 1 and then again in secondary level 2 [lower and upper secondary level]."* (Participant D - 13.06.22, Berlin)

To guarantee influence on the framework curriculum, its creators should be identified. Commission employees are often in the environment of CS teachers. Contact can be established, and awareness can be raised through teacher trainings.

Leverage connections within the teaching community to influence curriculum changes

AI should be included in the framework curriculum. Possibly as an elective and then as a compulsory subject or topic. The number of blocks should be determined. Clear definitions of competences and goals enable teachers to work more efficiently, which also has a positive effect on the students.

AI should be incorporated into the framework curriculum: define modules and competences

There should be more public relations work for the topics. Contact persons for this could be, for example, the Gesellschaft für Informatik (GI) and the Kultusministerkonferenz (KMK).

Amplify the importance of AI in education

AI can be anchored in specialised seminars for teachers. Concrete exercises of application can be developed in further training such as TrainDL. Other possible formats include familiarisation days with exercises from the TrainDL training. To test the teaching, various implementation formats can be considered, such as teaching it after the grading period, as trial blocks, or as a class workshop. These can be carried out independently of regulations, as no institutional resistance is expected.

AI can be introduced to teachers through specialised seminars and further training like TrainDL

Teacher education

Caution is needed when integrating the topics into teacher qualification: *"[...] you have to see where this could realistically happen because there are very often demands to do this and that as part of teacher formation, and I see a danger in overburdening teacher education. Creating a hodgepodge where each topic is covered only once in a seminar unit would not do justice to many of them."* (Participant B - 13.06.22, Berlin)

There's a risk of overloading teacher training with too many topics

Nevertheless, there are convincing arguments for including these topics in teacher's development: DL and AI are modern fields in transition, which have an increasing importance in the digital world. The integration of these topics can therefore be seen as a measure in the context of promoting and demanding digitalisation among teachers.

Arguments for integration: DL and AI are evolving fields crucial to the digital age

DL should be a part of teacher education as there are new tools like Orange3 that are easier to use compared to other programmes

Opinions on inclusion of DL vary

like SPSS. Orange3 focuses more on content aspects and allows for more efficient knowledge transfer and easier exchange with students. However, there are also arguments against the inclusion of DL in education. The right approach for integration into teacher education and the framework curriculum is still missing. Moreover, teachers should already have mastered this topic, as DL is already integrated via the database section of the framework curriculum.

Integration of AI makes sense

The integration of AI into teacher education is reasonable because AI is not sufficiently represented in tertiary/university education, and it has many points of intersection with other subject areas. Including AI in teacher qualification would be a logical consequence of curriculum integration: *"That implies that if it is part of the curriculum and if I believe that it belongs to the curriculum, then I naturally also believe that it should be part of teacher qualification, with the same requirements for competency development as I have for the curriculum framework."* (Participant F - 13.06.22, Berlin)

School curriculum

Multiple avenues for DL and AI introduction in classrooms exist

There are various possibilities for integrating DL and AI into the classroom. One possibility is to integrate the topics into the school curriculum via broader contexts such as ethics and media. If schools are given the freedom to choose their own topics, there should be no major problems.

Teacher's autonomy takes priority over school curriculum.

The school curriculum leaves little room for computer science teaching and is more focused on technical implementation and enabling teaching. Instead of strictly adhering to the school curriculum, the responsibility often lies with the teachers. Centralizing the inclusion of topics within the computer science framework curriculum is

advisable: *"And I also don't understand why every school has to map this individually and why something like this is not embedded centrally in the subject curricula, but at every school the work is done anew to see what fits where."* (Participant B - 13.06.22, Berlin)

Integration of DL and Orange3 is not convincing

Regarding DL, colleagues in CS would have to be convinced to integrate it, especially with Orange3. Otherwise, no major problems expected.

Integration of AI requires high-level commitment and uniform knowledge level from the teachers

The integration of AI is possible, but problems may arise under certain conditions. The implementation of AI within the curriculum requires a high level of commitment on the part of the teachers, agreements within the subject department, the assessment of the students' level of knowledge and the alignment with the curriculum. Another prerequisite is an equal level of knowledge among the teaching staff: *"They all have to have done at least this training [TrainDL] together, which I did, so that we can teach similarly, because I don't know if I can teach other teachers so quickly what I'm doing."* (Participant A - 13.06.22, Berlin)

Reprioritisation of the curriculum or guidelines might be needed

A reprioritisation of the curriculum or guidelines from the school management might be necessary.

Teachers typically exhibit reluctance toward integrating new subjects like DL and AI, perceiving them as added burdens

Teachers are generally assessed as not having a positive attitude towards the school curriculum or the integration of new topics such as DL and AI into it. Because the elaboration and writing down of new concepts is often perceived as disruptive, as it is not part of everyday professional life. In addition, there has been confusion in the past about the terms and concepts written down in the framework curric-

ulum. However, most teachers are motivated to implement new impulses and report positively after training events such as TrainDL and pass on what they have learned to their colleagues. Trainings like this are perceived as a service, while the curriculum is seen more as a task or a burden.

Challenges of the diverse student backgrounds

The heterogeneous student body poses a challenge when it comes to covering AI content with all classes.

Barriers

Only few Institutional barriers are expected

Few institutional barriers are expected as institutional awareness of DL and AI is anticipated: *"Well, most of the time you do find a way, if you have good material, good examples, so I see few institutional barriers now."* (Participant A - 13.06.22, Berlin)

Shifting priorities towards specific subjects may result in other subjects receiving less attention or resources

Regarding the acquisition of equipment of what is necessary for implementation, hardly any barriers are expected. Reprioritisation in favour of certain subjects can lead to other subjects being disadvantaged.

Unplugged exercises are possible, while Orange3 is problematic

There are both challenges and potential for successful integration when implementing plugged exercises with Orange3. The initial obstacle might be the installation process on computers by the system administrator. Yet, even with less-than-ideal equipment, executing plugged exercises in the classroom remains feasible.

Integration needs concept or a strategy with the subject leader

The development of a joint strategy with the subject leaders of computer science is a prerequisite. Unplugged exercises can be implemented without any problems, as no PCs are needed. However, this requires an implementable concept or a joint strategy to gain

acceptance. If there is no acceptance, the AI could also be imparted alternatively under the topic "Algorithm".

Teacher workload is a challenge

The workload of teachers poses an obstacle to a high-quality teaching and school development overall. Many ideas cannot be implemented or reflected upon due to the prioritization of daily tasks.

While schools have the responsibility to prepare children for the future, this objective is perceived differently in practice: *"When you see how our society is developing, you have to say that school is miles behind in many aspects, although it should prepare the youngest for the future. Ideally, schools should be at the forefront of shaping the future, but in my experience, it seems to be rather the opposite."* (Participant B - 13.06.22, Berlin)

D. *(Possible) changes through the integration of DL and AI in the framework curricula:*

Students

Integrating DL and AI allows students to reflect on the applications and devices they use daily

In the context of DL and AI, reflection on the applications and devices of daily use becomes possible. This allows to understand the functionality of platforms like Instagram and smartphone apps by recognizing that the display of content on them is purposeful. The ability to reflect leads to a critical understanding and promotes maturity in dealing with technology. Otherwise, there is a risk of vulnerability: *"Otherwise, students simply receive information passively and accept it without question, leaving them vulnerable to whatever they're presented with."* (Participant A - 13.06.22, Berlin)

Social implications of DL and AI are important

There is an opportunity to raise awareness about DL and AI. Here, the societal aspects are more important than the purely technical

aspects, whereby understanding the technical basics is a prerequisite for understanding the societal dimension.

In the context of AI, it is about promoting interest and reflection, because the topic is also applied in other subjects like biology and mathematics. A targeted demystification of devices and topics in connection with AI can take place.

Fostering interest for, and demystification of AI is important

Furthermore, competences can be developed to differentiate and evaluate real and fake news. This serves as a basis for obtaining and assessing trustworthy information. Without these competences, a meaningful use of information is not possible. It is also important to learn about the weaknesses of AI. This can thus be better understood, handled, and critically assessed.

Learning about weaknesses of AI is needed to evaluate information sources

Regarding AI, a professionalisation of teaching is necessary to be able to convey complex topics such as neural networks. In addition, the lessons can then serve to reach parents through their children and familiarise them with the topic of AI.

Professionalisation of teaching is necessary to convey complex topics

School authorities and schools

School authorities would have to think about the resources needed to support integration. Resources refer to both teaching materials and the provision of in-service training for teachers.

Providing teaching material and training is needed

The teaching of AI in CS is becoming increasingly important: *"I could imagine that the relevance of the topic would be recognised by saying that maybe not only computer science teachers should be knowledgeable about it, but that it should be a topic in other subjects as well."* (Participant

More AI trainings, also for non-CS teachers is needed

B - 13.06.22, Berlin). In addition, the support for further trainings in the field of AI could be expanded.

School image could motivate introducing AI

Image plays a role for both school authorities and schools. School authorities would demonstrate their modernity and relevance by dealing with contemporary issues such as AI: *"So, when I engage with such modern topics, it already appears more appealing than dealing with old-fashioned ones like what's the name of the keyboard, what's the name of the mouse?"* (Participant A - 13.06.22, Berlin)

Schools that offer AI in the classroom are likely to have a more attractive image than those that do not. Thus, in the future, the attractiveness of a school might also depend on the relevance and integration of AI.

Society

Modern algorithms can have significant societal implications

Current algorithms have disruptive potential: *"Yes, one could almost take this in a philosophical direction, where AI and algorithms on Facebook determine politics and partially drive and accelerate societal divisions by implementing the algorithms as they were designed."* (Participant A - 13.06.22, Berlin) However, if we understand how DL and AI work, these societal divisions can be prevented, and the protection of democracy can be promoted.

AI literacy can foster democracy

AI enables more competent decision-making, including in terms of better opinion formation, for example regarding the right to demonstrate.

DL and AI competences can inform media consumption

DL and AI encourage reflection on how, for example, personalised selection of services based on tracked user behaviour works.

Knowledge about these topics enables a critical examination of media consumption and its selection.

DL and AI competences can foster broader public discourse eon AI

The discussion about these topics in society needs to be promoted, both about their positive and negative aspects. Currently, this discourse is perceived as follows: *"I sometimes have the feeling that it is only discussed in certain circles and not in the broader society because it is such a big black box that cannot be looked into."* (Participant D - 13.06.22, Berlin)

We should aim for a broader discussion about the role of AI in the society: *"The question always remains, what do we want to have as AI and what do we do not want to have? And I think that is a very relevant question, and my wish is that as many people as possible can competently assess what it means to have a task carried out by AI."* (Participant B - 13.06.22, Berlin)

Making AI compulsory in framework curriculum to incentivize teachers for further training

Regarding teachers, integrating AI as a compulsory component in the framework curriculum could lead to more teachers taking part in corresponding further trainings and thus developing an increased interest in the topic.

E. Training Feedback and Potential for Improvement:

Length

7 hours are too short

The trainings are optimal and well-designed, but they could be a bit longer and expanded, as the current length is almost too short. In the given length, however, there was neither over- nor under-challenging and there was no boredom.

Longer trainings and more content is desired

The training sessions are well filled with the content covered and it would not be problematic if the trainings were longer due to additional topics.

7 hours per day might be a limit for some

There are advantages and disadvantages to different training formats. A one-hour extension (within the same day) of the TrainDL training could be a problem for some: *"I'm not sure now if an extra hour would have overwhelmed me already, I probably would have forgotten the first part by then."* (Participant A - 13.06.22, Berlin)

Full-day training is well received, yet might require taking a day off

Full-day events have the advantage that they sometimes offer better motivation and participation opportunities than half-day or afternoon events: *"Most of the time, the trainings are offered in the afternoon, and I always find that challenging. It's too short to really get into it and then when you have a day when you're not present at school, then you have the space."* (Participant B - 13.06.22, Berlin) However, not all teachers can take a whole day off to attend a full-day training.

Afternoon events might work for beginners, yet teaching workload is a problem

Afternoon events, on the other hand, can serve as an introduction for teachers who are not yet intensively familiar with the topic. However, the disadvantage is that the preparation for the next teaching day is postponed to the evening.

Therefore, it is suggested to offer full day as well as half day or afternoon in-service trainings to better meet the needs of all teachers.

Content

Content received positive feedback

Participants were positive and praised the inclusion of digital content such as Orange3 in the training. The "Beat the Robot" game was particularly praised because it was motivating to be able to defeat the robot eventually. In addition, the participants found the examples presented at the beginning easy to follow. They also liked the idea of unplugged exercises.

Criticism: no content/material for intermediate level, no clear thread for a series of lessons

There was criticism about a lack of content or material for the intermediate level. The training seemed to be geared more towards beginners and advanced learners and it was unclear if this was intended. Ideas and materials for the middle of a lesson series were missed: *"For example, what does the fourth lesson on AI look like? I can imagine the first lesson well, also the tenth lesson, but somehow in terms of content, what do I do in the fourth lesson?"* (Participant A - 13.06.22, Berlin). Furthermore, the participants perceived the topics of DL and AI as very different, lacking a connecting thread. There was also criticism that the topics covered were not sufficiently in-depth.

DL/Orange3 content was challenging

The AI topics in the morning were still understandable, but as the content became technically more advanced (DL, Orange3), it posed challenges for the participants,

"Beat the robot" and guessing game needs more time

With the "Beat the Robot" game, the participants would have needed more time to define the rules, otherwise the connection would be lost in the exercise. The guessing game with pictures of people at the very beginning of the training was criticised for not being obvious enough.

Wish for further AI topic, no DL

The participants expressed the wish for certain contents. In the field of AI, topics such as fake news, image manipulation and recognition, speech recognition and translation, machine learning with Python, robotics (especially in relation to human-like robots), Amazon Go and social and ethical aspects were mentioned. Some of these topics would obviously require more training time and could lead to further topics.

Suggestion for web-based games, instead of analogue games

In addition, a web application for games was mentioned, which was presented at the Computer Science Day ("Informatiktag"). This is interesting for teachers who do not want to do analogue games with their students.

Format

Face-to-face format worked well

The implementation of the training as a face-to-face event has met with approval.

Interaction

More frontal teaching is desired

The proportion of the frontal teaching was perceived as slightly too small. Theoretical introductions were praised: *"I also found these input lectures were really super to the point, very helpful to classify what then happened in the application phase."* (Participant B - 13.06.22, Berlin)

Good mix between theory and exercises was praised

The ratio was found to be both good and optimal. It was pleasing that there was no pure frontal teaching, but that the participants were actively involved in the activities.

Do-it-yourself approach and group work were praised

The exercises enabled the participants to try things out for themselves, which contributed to reflection on the use in the classroom. Working together in groups of two also encouraged reflection.

TrainDL Training (other factors, overall view)

Structure of training was well-received

The structure of the training was perceived as clear and the build-up from rule-based learning to supervised learning was seen as successful. This structure also corresponds to the design of classroom instruction, as it presents an argumentative arc.

Providing material as a ZIP file or in other formats that allow sharing with colleagues is helpful

The materials on AI are well suited for both introductory and advanced purposes. The provision of the materials as take-away copy templates was particularly praised. The ability to provide the materials as a zip file has several advantages: *"I liked the fact that all the material is online in this ZIP file, you don't have to laboriously put it together, but everything that was covered today is in one place. Then you have it at your disposal. I'll have another look at some of it."* (Participant B - 13.06.22, Berlin) Sharing with colleagues, for example in the form of a printed hardcover book, would also be possible with this. However, the wish was expressed to also be able to purchase the materials as a class or school set in the form of printed boards.

Splitting the course to use a station approach was suggested

Splitting the course after the lunch break would be an option: *"What one could think about is whether to split the course after the lunch break and ask who is interested in learning more about this and who would prefer to deepen and further apply this at the level of the topics worked on in the morning? Who found the morning a bit too fast and would like to apply it a bit more in peace and quiet? And for those who say, I have that on my radar now, and I want to continue, that would be an option."* (Participant B - 13.06.22, Berlin)

The training was found to be successful and interesting. The training would be recommended to others. The pleasant atmosphere during the training contributed to its enjoyment. The efficiency of the training was considered very high, with little room for improvement. The pace of the training was felt to be appropriate, and the breaks were well planned.

Training received overall positive feedback

F. Wishes for education policy

In the area of school conditions, it is proposed to introduce smaller classes and make hardware upgrades based on needs.

Smaller classes and hardware upgrades can be introduced

AI should be included in school textbooks.

AI should be included into textbooks

It is stressed that the existence of subjects like DL and AI should be acknowledged.

Acknowledgement of DL and AI at the policy level is necessary

In addition, AI should be taken more seriously, and a debate should be held that takes into account positive and negative implications because currently the discourse is very polarised: *"Either it is totally positive because there is something new, mostly in the medical field. And on the other hand, it is demonised when it comes to social media. It would be better if it were more balanced in that regard."* (Participant D - 13.06.22, Berlin)

More balanced out debate is needed

An integration of AI in the framework curricula is desired, whereby it could be oriented towards other federal states such as for example Bavaria.

Integration in framework curriculum

Integration of AI in the subject of computer science could increase the attractiveness of the subject in lower secondary school (grade level 5-10). Good, unplugged materials are important for this, so that simple ideas can be taught in a playful way. In addition, it is suggested to integrate AI either as a compulsory subject in computer science or as an independent subject. The integration could be age-specific (taught once) or implemented in a spiral curriculum manner (taught continuously), with a suggestion of starting from 7th/8th grade (lower secondary level).

AI integration is possible starting from 5th or 7th grade

CS should be offered as a compulsory subject: *"[...] I would like to see compulsory teaching in computer science, a compulsory subject, because this is the only way to ensure that all students get to know the contents that we are working on in computer science lessons. If students choose computer science, those who are interested in these topics anyway will also learn something about AI and everyone else will not."* (Participant B - 13.06.22, Berlin)

CS should be offered as a compulsory subject

As CS is currently not a compulsory subject, teachers are putting too much work into an offer that a large part of the students are currently not taking advantage of. Ideally, integration of AI should already take place in lower secondary level (grade level 5-10).

Integration can start from grades 5-10

The creation of further in-service training, such as TrainDL, should be sought out. On the part of the teachers, sufficient time resources should be secured to be able to attend these in-service trainings.

More trainings similar to TrainDL are needed. Teacher need time to attend such events.

Additionally, there's a call for hands-on political figures to deliver information on fake news. This might take the shape of school informational sessions following in-class introductions (like those about

The subject of fake news requires attention

AI). Contacts should be easily accessible to share insights when needed.

G. Other Themes

There is a need for further training

A mere integration in the framework curriculum does not result in automatic guarantee of teacher proficiency, there is a need for more training: *"Just because dealing with AI is included in the curriculum doesn't mean that everyone will be able to do it after a session like today."* (Participant A - 13.06.22, Berlin)

Grade level(s) should be clearly defined

There is also a difficulty in deciding in which grade level the topics should be implemented. The current framework curriculum in Berlin dates from 2006 and will soon be replaced by a new one. It is also noted that the framework curriculum is kept very concise and therefore offers possibilities for action [note: it remains unclear here, however, whether only the subject of CS is meant].

The evaluation survey question on the curriculum is confusing

AI is not fully integrated into the curriculum in Berlin. It is not mandatory in CS, but it is sometimes addressed in other subjects like English (semester topic on Science and Technology). The question about the integration of the framework curriculum in relation to AI is unclear in the evaluation questionnaire, as AI is already integrated, but only as an optional specialisation area.

Training series could be a suitable format

One suggestion is to extend the TrainDL training and integrate it into a training series, for example with aspects of programming with TensorFlow or Python, which would be interesting for use at the secondary level. These training courses motivate the participants to work intensively on the topics.

Difficulties with AI definition

After previous trainings and discussions, it was unclear how exactly to define the concept. Only after the TrainDL training it became clear that AI does not have a clear one. Because of these different definitions of the topic, there was sometimes a challenge in the competency test to decide what the right solutions are. AI is a loaded buzzword that is present everywhere and becoming increasingly visible, as modern computational capacities enable problem-solving associated with AI.

2.3.2 Follow-up interviews

A. Training:

Overall training was praised

In retrospect, the training was rated as very successful and constructive. It has succeeded in reinforcing the importance of the topic of DL and AI. The mix of new as well as familiar content as well as the breadth of content was rated positively. he "Shopping Cart" exercise received special mention. The material from the training is suitable for teaching, particularly in segments where there were no pre-existing materials.

Orange3 is challenging

Orange3 exercise was praised as a good, new, open-source-based, albeit conceptually demanding tool. The program's usability issues were seen as challenging: *“And when you work on it yourself, it's also a bit awkward, as is usual in computer science. And then you're kind of left hanging in the air.”* (Participant A - Follow-up to 13.06.22, Berlin) This represents a hurdle or uncertainty for the possible application in lessons and further work with this tool.

In-depth AI-examples are needed

Both interviewees have continued to work on the topic of AI, but were not able to provide in-depth examples.

B. Integration (after training):

The level of integration of DL and AI into the teaching between the training session and the follow-up interview can be divided into three areas: (1) integrated, (2) not integrated, and (3) uncertain or (vaguely) planned integration for the future.

(Integrated) Two courses on databases were taught by one teacher. However, there is uncertainty about whether DL was addressed at all or, if it was, whether it was only implicitly included.

Not clear if DL is part of teaching databases

Because, in contrast to AI, a person could not recall the training contents on DL: *"I have to admit, after half a year, I was also interested in what there is to say about data literacy. And I can't remember much about that. About AI, of course. [...] And at that time I was also interested in what is meant by data literacy. What could classes on data literacy look like? And I have little memory of that now."* (Participant B - Follow-up to 13.06.22, Berlin)

Not clear what DL means

(Not integrated) However, no proactive instruction on AI was integrated, even though it had been done in the past prior to the training. On the one hand, this was explained by the analogue approach of the exercises.

AI was not integrated because the exercises were not digital

"The [analogue exercises] are great, but the students come to computer science classes because they work with computers. That's what makes the CS classes different. And if you then get a method like that again [analogue], which they also have in another subject, then they are not enthusiastic." (Participant B - Follow-up to 13.06.22, Berlin)

Students want to have digital exercises

AI did not fit into the planned CS course topics

On the other hand, the integration did not happen, because the topic of AI did not fit into the planned range of the CS topics. In addition, (Orange3) was criticised for lacking the actual programming experience. *"And if I were to do it in Sek 2 [11-13 grade level], I wouldn't do it with Orange, because I have the feeling that it's too far away for advanced students. They should program it. They should be able to do that. They should master it."* (Participant A - Follow-up to 13.06.22, Berlin)

Hesitation to integrate AI due to unsuccessful experience in the past

(Uncertain) There is uncertainty about implementing AI topics/materials in class, due to previous unsuccessful experiences where the integration did not go optimally. However, the conviction to want to implement it still exists to some extent: *"I would like to do it again. [...] Because it plays an increasingly important role. In all of our lives. [...] The goal is to somehow teach the students understanding so that they can contribute later on. [...] Ideally, they should be able to make meaningful decisions about the use of AI."* (Participant B Follow-up to 13.06.22, Berlin)

C. Training feedback and potential for improvement:

The topic of LLM is very relevant

Large language models (LLM) are generally seen as an increasingly relevant topic.

ChatGPT has a potential to disrupt didactical approaches

For teachers, a possible full automation associated with the LLM tools represents a change in the expectations placed on the students. *"Basically, we don't need to assign homework any more where students have to answer something. ChatGPT does that for them. You can give homework like learning vocabulary or something like that. ChatGPT also does programming. It's all kind of pointless."* (Participant A - Follow-up to 13.06.22, Berlin)

CS needs to be introduced as a compulsory subject

Another aspect of subject integration concerns making computer science a compulsory subject: *"But if the aim is to educate students in such a way that they can later, make well-reflected decisions about the use of AI, the first step would indeed be to reach all students in the first place."* (Participant B - Follow-up to 13.06.22, Berlin). It is paramount for students to be able to program, i.e., work with computers.

More advanced AI training is needed

For teacher training in general and for TrainDL specifically, there is a recognized need to go beyond a descriptive level of AI and engage with the topic at a deeper level. As part of it, it is suggested to further explore the feasibility of training content for both lower and upper secondary level.

Material should explain complexity/black-box and be prepared for a specified grade levels

The suggestion for the material was to emphasize explaining its complexity, ensuring teachers don't view it as a black-box concept. It was also expressed a need for complete lesson plans for concrete grade levels with a specified previous knowledge.

Full-day trainings are preferred over a series of events

Further trainings (if possible, also within the framework of TrainDL) are welcomed. The interviewees preferred full-day (follow-up) trainings (with optional afternoon trainings for a deeper dive) to a series of trainings. The latter is seen critically, as it requires attendance of each session within the series.

More teacher training with a focus son AI is needed

In the medium term, there's a growing desire to enhance teacher training programs with an emphasis on AI, which is justified by the relevance of the topic: *"With the importance that AI has today, I think every computer science teacher should be able to provide information and address it themselves."* (Participant B - Follow-Up on 13.06.22, Berlin)

2.4 Summary

Based on the pre-survey conducted immediately prior to the training, which was completed by 23 out of 24 participants, the CS training group in Berlin primarily consisted of in-service teachers, with only five participants in the final stage of their training. The participants had a variety of second subjects, with mathematics and physics being the most prevalent. It was observed that most of the participants had previous experience in teaching DL in their classes. However, the training exhibited a relatively low representation of women, which reflects the existing gender ratio among CS teachers in Germany.

The findings from the evaluation survey (albeit based on a small sample size) indicate that on average, the training had a positive impact on teachers' perceived competences on how to use DL and AI in their teaching. For both topics DL and AI, participants reported an increase in perceived competences in how to use the content in class. In addition, participants demonstrated a generally high willingness to invest time and effort in incorporating DL and AI into their teaching, with a stronger inclination observed for AI compared to DL. Participants expressed a moderate level of agreement regarding whether they have acquired enough competences to teach the learned content in their classrooms.

Both self-report and objective knowledge questions provide evidence of some improvement in understanding of AI concepts. However, while participants on average reported an improvement in the understanding of DL concepts, the objective knowledge test could not confirm that increase: after the training, participants demonstrated a higher variability in their answers to the DL knowledge questions, with some participants showing a higher level of knowledge and others demonstrating a lower level.

The post-evaluation survey yielded additional insights into the perceived challenges associated with the DL content of the training. While most participants regarded topics of classical AI and machine learning (ML) as highly suitable for their teaching, they found the topic of data lifecycle to be less suitable. Furthermore, the data lifecycle exercises involving the Orange3 tool received lower ratings compared to the other exercises. These findings indicate that participants consider classical AI and ML topics to be well-suited for teaching, but improvements are needed in the delivery and effectiveness of the data lifecycle topic and the exercise using the Orange3 software.

Out of the seven participants who completed the follow-up survey approximately six months after the training, only three reported incorporating DL and AI content into their classes. Among these participants, two had already been teaching both topics prior to the training. Interestingly, all four respondents who had not started teaching DL or AI at the time of the follow-up survey expressed their intention to do so in the future. The planned topics for future instruction mainly revolve around AI, with no participants indicating plans to teach the topic of "data lifecycle and the basics of statistical data analysis and interpretation" (e.g., Orange3).

For both DL and AI, training participants have a slight inclination towards the opinion that these topics are lacking in the current CS framework curriculum. This perception remains consistent and does not change immediately after the training. The qualitative interviews shed light on the results. On the one hand, the survey question was confusing, as AI is already integrated, but only as an optional specialisation area. On the other hand, while integration is generally viewed positively, there is a challenge of removing content from the existing framework curriculum.

After the workshop, participants expressed a higher belief that teaching AI will provide added value to students compared to their pre-training scores. The scores for the similar item on DL showed no significant difference between the pre- and post-training, with both scores falling in the moderate range. The latter is consistent with

other results stressing the problem of the DL content. According to the post-survey, there was a strong consensus among participants regarding the societal importance of DL and AI. Additionally, the post-survey results suggest that teachers hold positive expectations about students' interest and their own ability to generate excitement and engagement for DL and AI topics, with the AI topic receiving higher ratings in this regard.

Overall, the quantitative findings demonstrate the positive impact of the 7-hour CS training in Berlin on enhancing teachers' perceived competences on how to teach AI content in class as well as their objective knowledge of the introduced AI concepts. However, the DL content proved to be more challenging. The results also indicate that participants express positive attitudes towards both topics, albeit it is uncertain whether the training directly influenced these attitudes, as they were measured only post-training. The moderate consensus among participants regarding the statement on having acquired sufficient competences to teach the learned content in class implies that the training is not fully sufficient for the integration of the DL and AI topics into teaching.

According to the qualitative data, while the training received positive feedback and participants showcased an enhanced understanding of AI, there were identified areas of improvement. Teachers sought foundational knowledge and exposure to practical AI examples. They expressed a desire for pre-prepared teaching resources and guidance on lesson planning for specified grade levels. While basic AI and DL knowledge was grasped, mastering them is still a challenge. The complexity of AI, especially machine learning and neural networks, was recognized. Aspects like image recognition software or decision trees in Orange3 were seen by some as "black box". The interviewees viewed the training as a good entry point, but need further trainings to engage with the topics in more depth. More practical AI examples/exercises are needed.

The TrainDL training helped teachers to better lessons, but implementation of the content may require additional adaptation. Some exercises might need modification

for classroom use, and while the creation and preparation of materials will be time-consuming, the training has offered a solid foundation for doing so. Teachers wish for more concrete lesson plans for specified grade levels. The latter can make it easier to directly integrate the content into the teaching. The Orang3 exercise was viewed as complex and requiring more engagement with. Some teachers mentioned that Orange3 is better suited for more advanced classes, while others view the graphical interface and the lack of programming in the software as a major reason for not using it with advanced students.

Teachers acknowledged the high relevance of DL and AI topics for students and society at large. The integration of DL and AI into the framework curriculum is seen as pivotal given their societal and political implications. Some suggest that CS should be a mandatory subject for all at the secondary level. While daily relevance of AI makes it essential for secondary education, it is not clear how DL might fit and should be taught and at which level. AI is viewed as more student-friendly compared to the scientific intricacies of DL. The challenge for integration of AI into the framework curriculum is in determining which content might need to be omitted to make room for AI. There might be a need to adjust and reprioritize the curriculum. Teachers might show hesitance to include new topics such as DL and AI, seeing them as additional challenges, given a high workload. The wish for AI's inclusion in Berlin's curriculum is urgent, as states like Baden-Württemberg are already ahead. Ensuring AI integration in the curriculum will need collaboration with those shaping the framework curriculum. More awareness and outreach are needed in this area.

The integration of DL and AI into the framework curriculum will not result in teacher proficiency, there is a need for more teacher training in this area. Integrating DL and AI into teacher education is also viewed positively, albeit the overall workload of students/teachers should be taken into account. Incentive structures for reaching out to teachers should be discussed. Interestingly, some teachers mentioned integrating DL and AI beyond CS subjects, which means that teachers with non-CS subjects might also need to be trained in these topics.

Given the intricate nature of the subjects and the pressing need for DL and AI integration, it's crucial to consider appropriate formats for teacher training. While some educators lean towards full-day training sessions, they acknowledge that 7 hours might not suffice to dive deeply into these topics. Structured follow-up training sessions, which are easily accessible and suit the schedules of in-service teachers, are essential.

Teachers raised apprehensions about the easy accessibility of Language Learning Models (LLMs), especially in relation to academic integrity. This highlights the importance of providing AI support both in instructional techniques and in daily classroom and school practices.

3. CS in-service training, Lithuania

3.1 Sample

In total, 21 participants from Vilnius (15 participants), Kaunas (4 participants), and Panevezys (2 participants) participated in the training. All of them took part in the pre-evaluation survey. This group of participants, whose socio-demographic data is depicted in Figure 15, comprised predominantly of in-service teachers, except for one pre-service teacher. On average, participants were 45 years old, with ages spanning from 26 to 69 years. Half of the participants (11 out of 21) fell within the 40–49 years age bracket. Comparatively, this demographic profile aligns with the national teacher workforce in Lithuania, where, according to the OECD data (2023b), the 40-49 and 50-59 age groups constitute 26% and 37% of the total respectively, while only 14% belong to the 30-39 age bracket. About 20% of teachers in Lithuania are 60 years or older, signifying an older average age compared to other OECD countries (OECD 2023b).

In terms of gender distribution, the training participants in Lithuania mirrored national trends, with women representing approximately 57% (12 out of 21 participants)

of the total number of participants. The overall share of women teaching at the lower and upper secondary levels in Lithuania, is one of the highest among the OECD countries, at 82.4% and 78.4% respectively (OECD 2023b). Data on the share of female CS teachers in Lithuania were not readily accessible.

Most of the participants in the Vilnius training teach students from fifth to 12th grade, with six teachers who in addition teach at the primary level. Half of them were employed in a gymnasium, and the other half in various other types of secondary schools. In Lithuania, informatics is a compulsory subject at the lower secondary level and an optional one at the upper secondary level (European Commission / EACEA / Eurydice).

All the participants were teaching or studying CS as a subject, with the following additional subjects: mathematics, economics, physics, art, Lithuanian language, technology, and other. According to the pre-survey, mathematics was the most common second subjects among the participants. Before the workshop, most participants had some experience teaching DL in class. As Figure 16 shows, while about half of the participants never or only once taught AI in class, only 2 out of 21 respondents reported never teaching DL or covering DL-related topics in class on just one occasion.

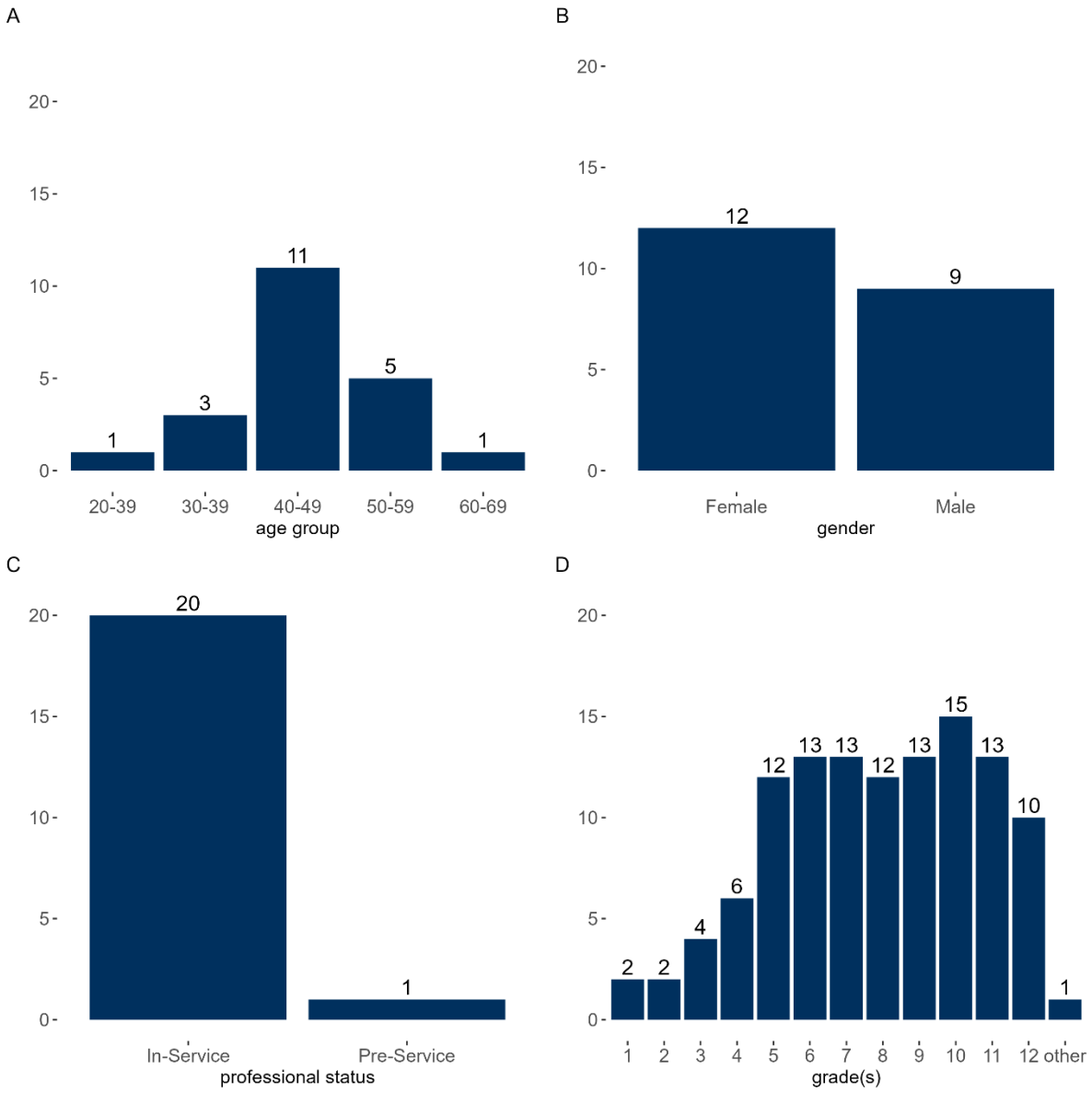


Figure 15 Socio-demographic data of the participants, 10.12.2022 Vilnius, pre-evaluation survey, n=21

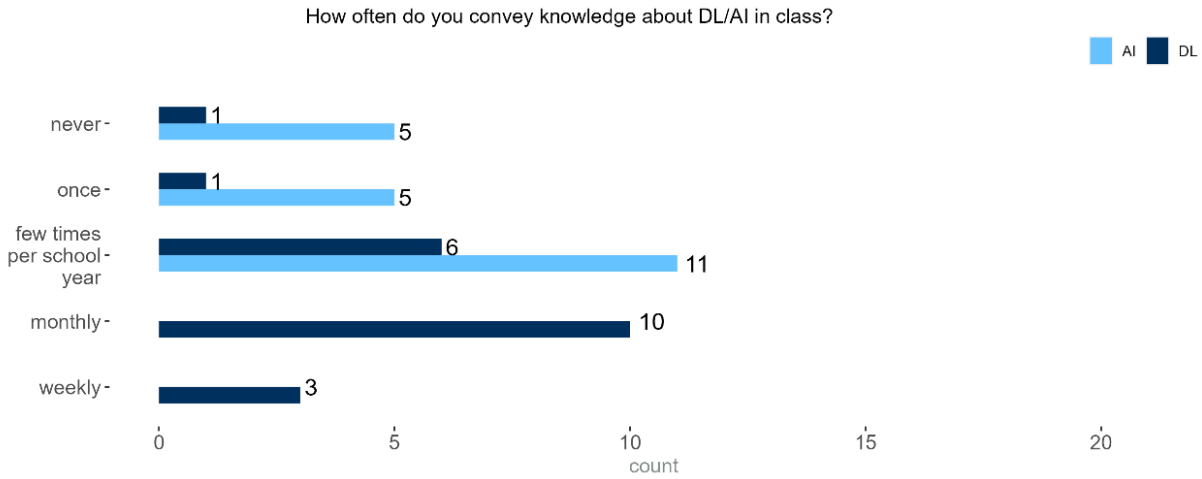


Figure 16 Experience with DL/AI, 10.12.2022 Vilnius, pre-evaluation survey, n=21

3.2 Quantitative results

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

As demonstrated in Figure 17 and Table 10, respondents reported an increased level of perceived competences in using DL and AI content in class following the training. For DL, the median score experienced a slight increase, rising from a high pre-training score of 4.5 to a post-training score of 5.0. The spread of the middle 50% of the data also slightly decreased for the post-training-values. The Wilcoxon signed-rank test did not indicate a statistically significant difference between the pre- and post-training DL values. In contrast, for AI, a substantial improvement was recorded. The median pre-survey score, initially lower at 3.0, surged to a post-training score of 5.0. The pre- and post-values were found to be different at the 1% statistical significance level, as determined by the Wilcoxon signed-rank test. The IQR slightly reduced for the post-training values from 2.0 to 1.25.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
I know how to use content about DL in the classroom	4.5	5.0	3.75	4.0	5.0	5.0	2	3	6	6
I know how to use content about AI in the classroom	3.0	5.0	2.0	3.75	4.0	5.0	1	3	5	6
I am willing to invest time and effort to incorporate DL into my teaching	-	5.0	-	5.0	-	6.0	-	3	-	6
I am willing to invest time and effort to incorporate AI into my teaching	-	5.5	-	5.0	-	6.0	-	3	-	6
After the training, I have gathered enough competences to teach the learned content in class	-	5.0	-	3.25	-	5.0	-	1		6

Table 10 Summary statistics of pre- and post-results for the survey items on teachers' perceived competences to use DL and AI content in class and post-survey results for additional items, 10.12.2022 Vilnius, pre-evaluation survey, n=16

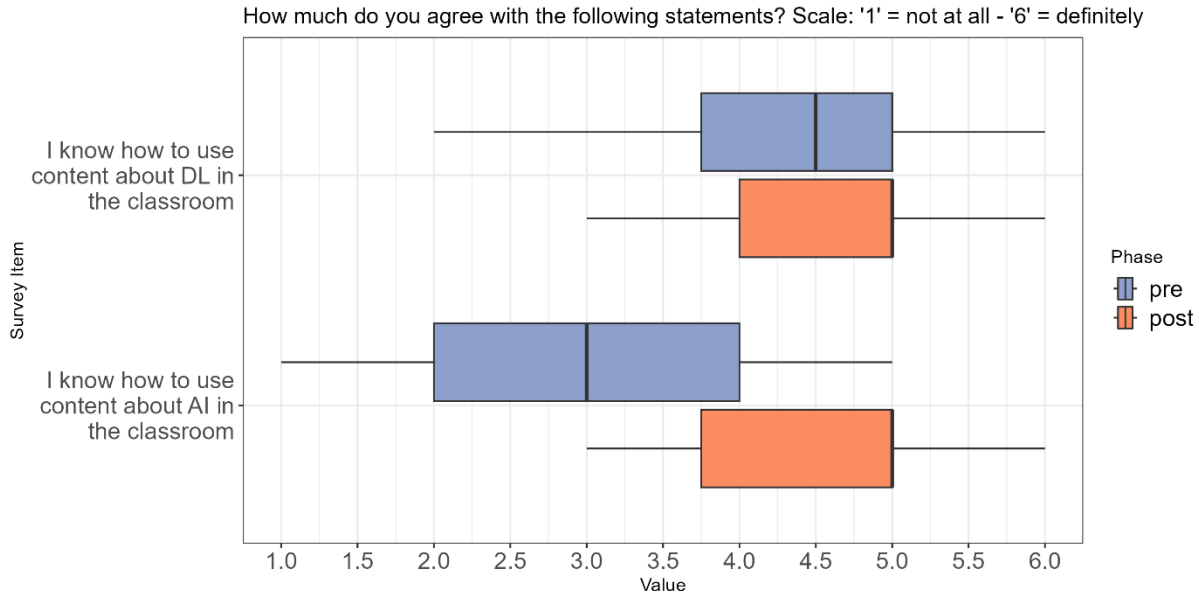


Figure 17 Boxplot comparison of pre- and post-results for the survey items on teachers' perceived competences to use DL and AI content in class, 10.12.2022 Vilnius, pre-evaluation survey, n=21

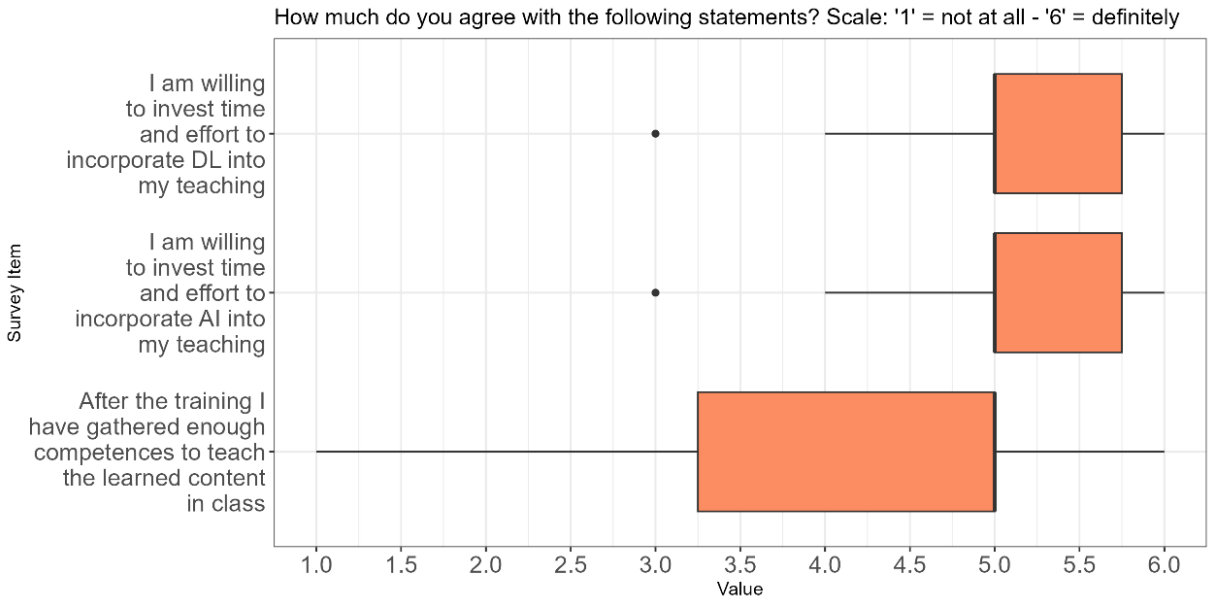


Figure 18 Boxplot of post results for the survey items on teachers' perceived competences and willingness to invest time and effort to incorporate DL and AI into their teaching, 10.12.2022 Vilnius, pre-evaluation survey, n=18

Figure 18 and Table 10 describe further survey items that were measured only in the post-version of the survey. Participants were asked to what extent they agree with the statements: "I am willing to invest time and effort to incorporate DL/AI into my teaching." For both DL and AI, most participants selected categories "5" and "6", where "6" represents "strongly agree." Results suggest that on average, participants expressed a strong willingness to invest time and effort to incorporate both DL and AI into their teaching. After the training, teachers were also asked to what extent they agree with the statement: "After the training, I have gathered enough competences to teach the learned content in class". Out of 18 respondents, 11 selected categories "5" or "6", while only 3 participants chose the categories "1" or "2". Overall, the statistics for this post-survey item indicate a leaning towards agreement that participants have gathered enough competences.

Overall, the findings indicate that the CS training enhanced teachers' perceived level of competences on how to use AI in their teaching. Although no significant differences were found for DL, the reported competences for the use of DL prior to the training were higher than those for AI. In addition, the post-survey measures revealed

a positive inclination among participants to dedicate time and effort towards integrating DL and AI into their teaching. On average, participants tend to agree, that they have gathered enough competences to teach the learned content in class.

3.2.2 Understanding of DL/AI topics presented in the training

When it comes to the self-assessment, for both DL and AI questions, respondents reported an increase in their self-reported knowledge (see Figure 19 and Table 11) from a median of 3.0 to a median of 4.0 with a scale ranging between 1 and 5. The spread of the central 50% of the data also shrunk. The Wilcoxon signed-rank test indicated a statistically significant difference between the pre- and post-values for both DL and AI at the 1% level.

Respondents also demonstrated a notably higher median score for the objective knowledge tests on AI compared to the pre-training results, with the median increasing from 2.5 to 5.5 (see Figure 20). However, the IQR of the post-training values slightly increased. The Wilcoxon signed-rank test indicated that the pre- and post-values for the AI knowledge questions differ significantly, with the post-training scores being statistically higher at the 1% level. The test showed no statistically significant difference for the DL knowledge test scores.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Self-assessment for DL	3.0	4.0	3.0	4.0	4.0	4.0	2	3	4	4
Self-assessment for AI	3.0	4.0	3.0	3.5	3.88	4.0	1	3	5	5
Knowledge test DL	1.3	1.4	0.62	0.78	1.75	1.85	0.0	0.0	2.8	2.2
Knowledge test AI	2.5	5.5	2.0	4.63	3.88	6.88	1.0	1.0	7.0	7.5

Table 11 Summary statistics of pre- and post-results for self-assessment and knowledge test, 10.12.2022 Vilnius, pre- and post-self-assessment and knowledge test, n=14

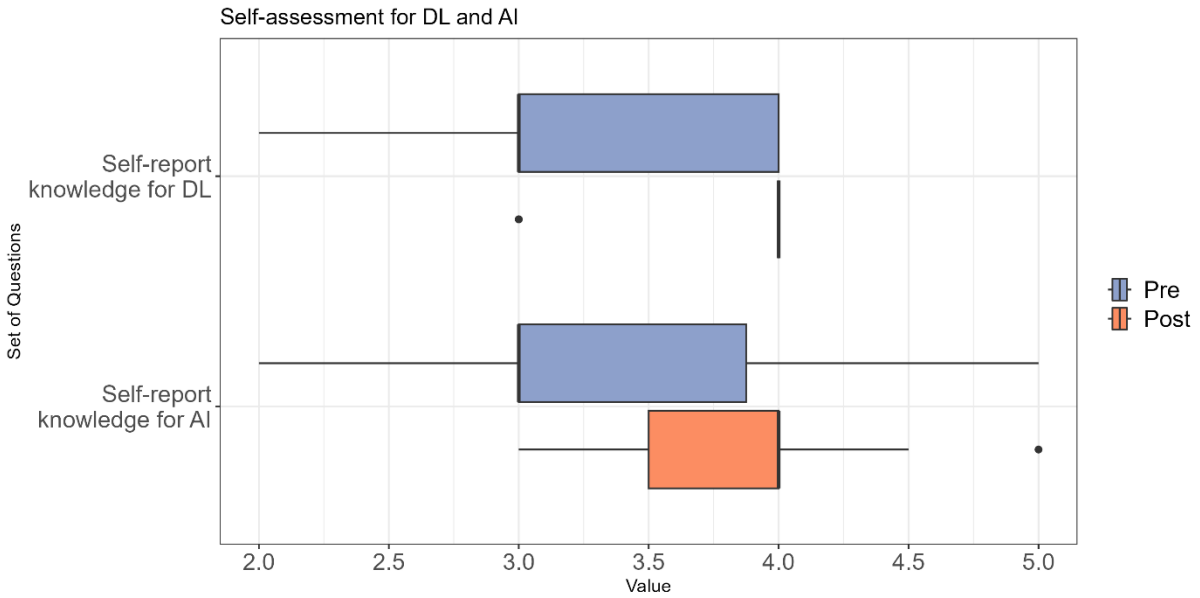


Figure 19 Boxplot comparison of pre- and post-results for self-report knowledge questions on DL (15 questions) and AI (10 questions), 10.12.2022 Vilnius, pre- and post-self-assessment and knowledge test, n=14

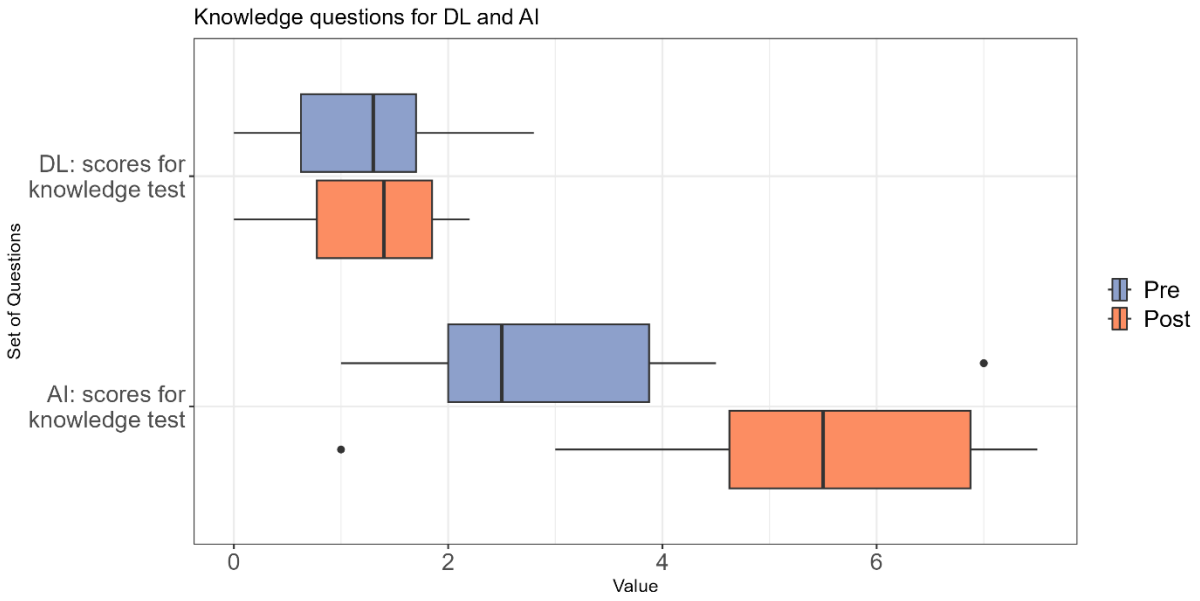


Figure 20 Boxplot comparison of pre- and post-results for knowledge questions on DL (4 questions) and AI (10 questions), 10.12.2022 Vilnius, pre- and post-self-assessment and knowledge test, n=14

3.2.3 Attitudes towards DL/AI

Table 12 and Figure 21 show the distribution for the survey items on teachers' perception of the importance of DL and AI, which were measured via both pre- and post-policy experimentation surveys: "In the future, teaching DL/AI will provide added value to students" and "I think the content of DL/AI is missing in the current framework curriculum."

For both DL and AI items, participants express a tendency to believe that DL/AI content is lacking in the current CS framework curriculum. According to the Wilcoxon signed-rank test, there is no statistically significant difference between the pre- and post-results. Interestingly, the IQR is higher after the training for both the DL and AI item. For the AI question, the post-training median score decreased from 5.5 to 5.0.

Similarly, there is no statistically significant difference for the pre- and post-results of the items: "In the future, teaching DL/AI will provide added value to students." However, for both items, respondents reported a relatively high score for the pre- as well as post-results, suggesting a high level of perceived importance of DL and AI for teaching already prior to the training. Similarly, to the curriculum item, the post-training median score for AI decreased from 6.0 to 5.0.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
In the future, teaching DL will provide added value to students	6.0	6.0	5.0	5.0	6.0	6.0	4	4	6	6
In the future, teaching AI will provide added value to students	6.0	5.0	5.0	5.0	6.0	6.0	3	3	6	6
I think the content of DL is missing in the current framework curriculum of computer science	4.5	5.0	4.0	4.0	5.25	6.0	1	3	6	6

I think the content of AI is missing in the current framework curriculum of computer science	5.5	5.0	5.0	4.0	6.0	6.0	4	3	6	6
The topic of DL is of enough societal importance to integrate it into the curriculum	-	5.5	-	4.25	-	6.0	-	2	-	6
The topic of AI is of enough societal importance to integrate it into the curriculum	-	5.0	-	5.0	-	6.0	-	3	-	6

Table 12 Summary statistics of pre- and post-results for a set of survey items on teachers' perception of importance of DL and AI, 10.12.2022 Vilnius, pre-evaluation survey, n=18

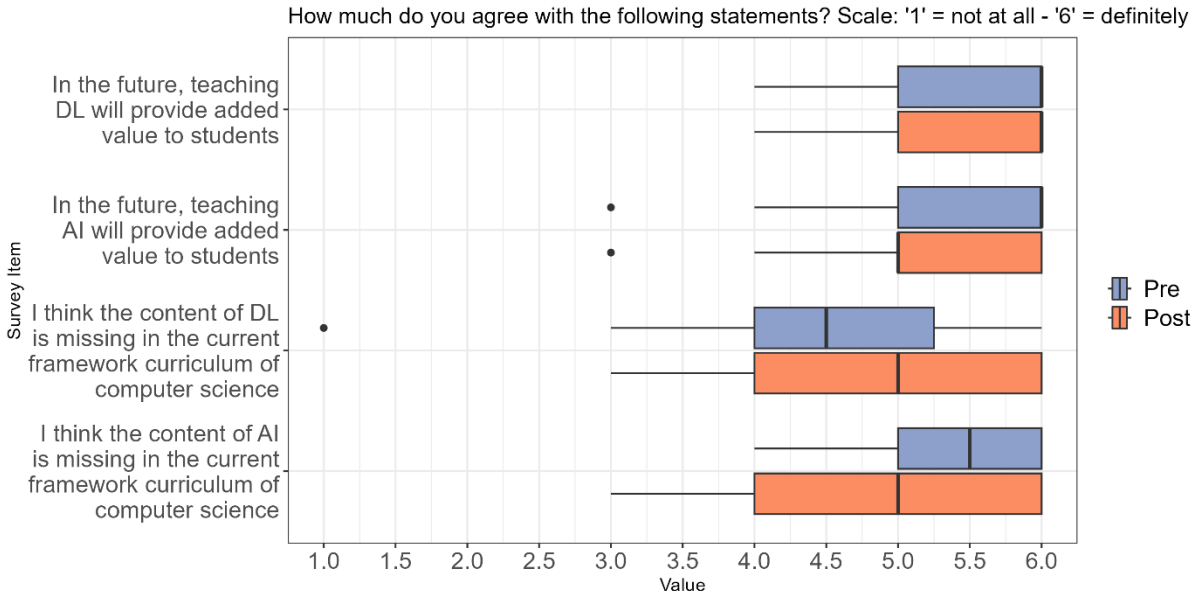


Figure 21 Boxplot comparison of pre- and post-results for a survey item on teachers' perception of importance of DL and AI, 10.12.2022 Vilnius, pre-evaluation survey, n=18

An additional survey item administered only in the post-survey revealed a strong consensus regarding the societal significance of DL and AI, with participants highly agreeing that these topics should be integrated into the curriculum (see Figure 22 and Table 12).

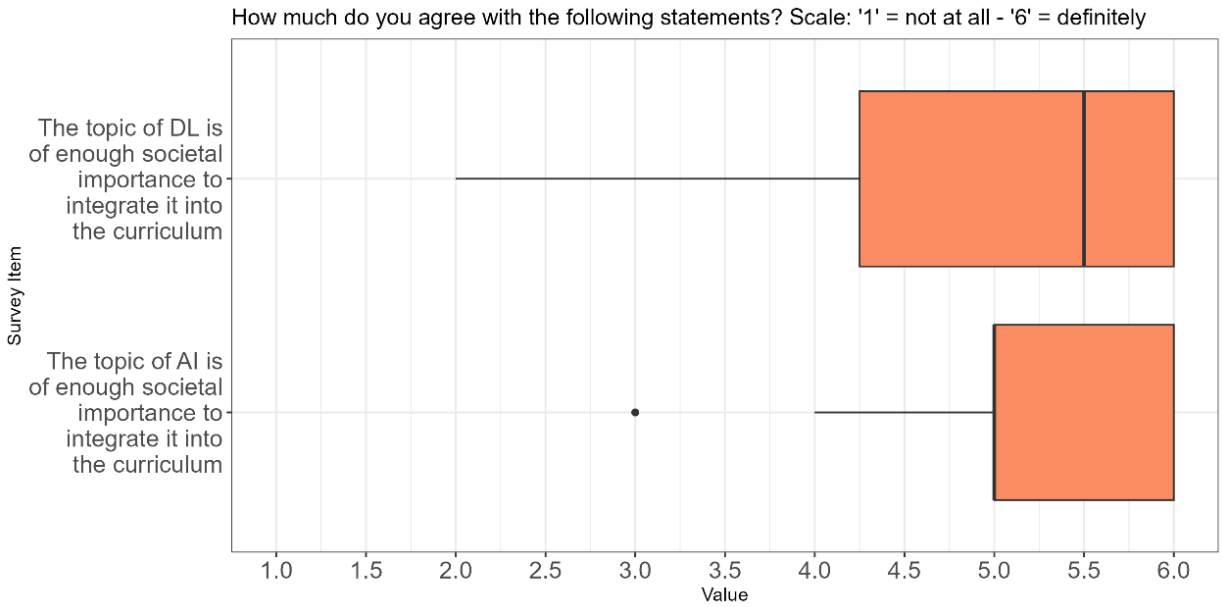


Figure 22 Boxplot comparison of post results for the survey items on teachers' perception of societal importance of DL and AI for integration into the curriculum, 10.12.2022 Vilnius, pre-evaluation survey, n=18

3.2.4 Perceptions of student engagement

In the post-survey, teachers were asked specific questions to assess their expectations regarding students' interest and their own perceived ability to generate student interest and engagement in DL and AI. These questions included statements such as "I can imagine my students show enthusiasm for the overall topic of DL/AI" and "I am confident that I can get students excited about DL/AI-projects".

The results depicted in Figure 23 and Table 13 demonstrate an overall positive expectation regarding students' interest, with a stronger emphasis on AI topics. The expected level of student enthusiasm for DL topics on average received a slightly lower rating compared to AI topics. This discrepancy implies that teachers anticipate a higher level of student engagement and excitement when it comes to AI compared to DL. Figure 24 and Table 13 also indicate a high level of confidence among teachers regarding their ability to motivate students for DL and AI projects.

Question	Median	Q1	Q3	Min.	Max.	Scale
I can imagine my students show enthusiasm for the overall topic of DL	4.0	4.0	5.0	3	6	1-6
I can imagine my students show enthusiasm for the overall topic of AI	5.0	4.0	6.0	3	6	1-6
I am confident that I can get students excited about DL-projects	3.0	3.0	3.0	2	4	1-4
I am confident that I can get students excited about AI-projects	3.0	3.0	3.75	2	4	1-4

Table 13 Summary statistics of post-survey results for items on perception of student engagement, 10.12.2022 Vilnius, pre-evaluation survey, n=18

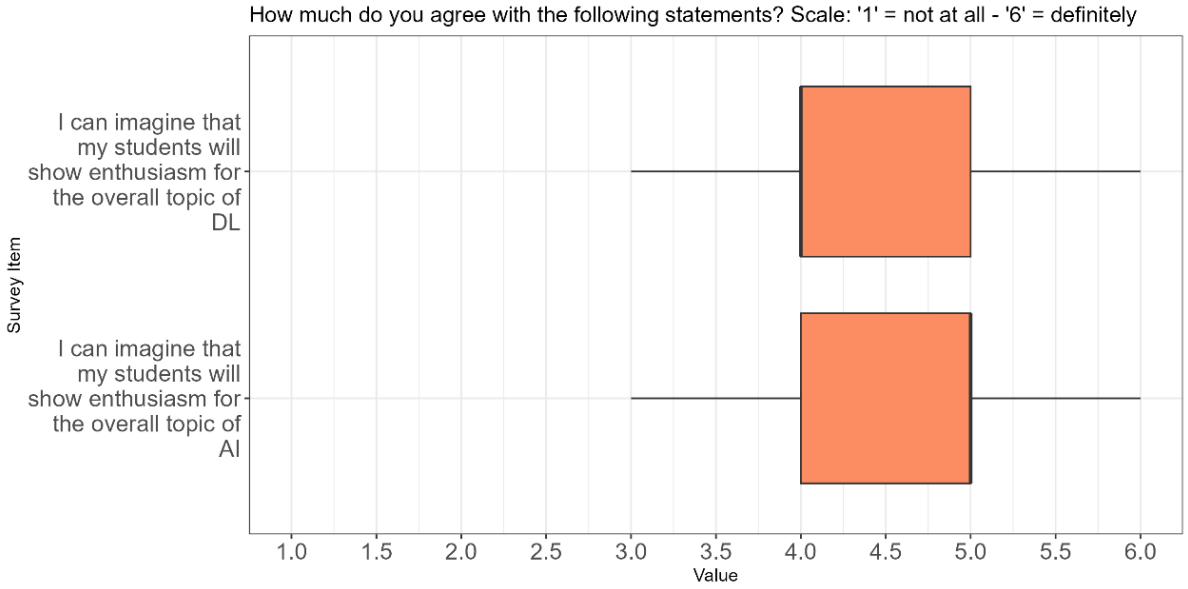


Figure 23 Boxplot of post results for the survey item "I can imagine that my students will show enthusiasm for the overall topic of DL/AI", 10.12.2022 Vilnius, pre-evaluation survey, n=18

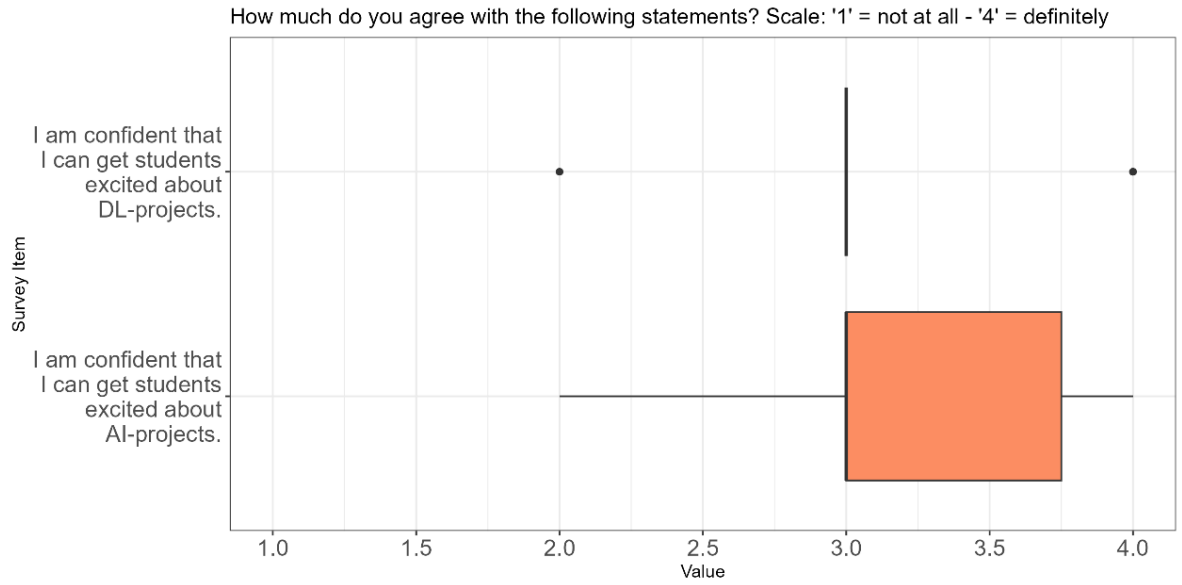


Figure 24 Boxplot of post results for the survey item "I am confident that I can get students excited about DL/AI-projects", 10.12.2022 Vilnius, pre-evaluation survey, n=18

3.2.5 Feedback on the learned content and format

The post-survey included a series of questions to assess participants' reactions to the suitability of the topics, practical materials and exercises presented in the training. Figure 25 and Figure 26 show that participants reported a very high level of perceived suitability of the topics and exercises. Approximately 83% of respondents selected a score of "5" or "6" (where "6" indicates "very well suited") for the topics of classical AI, 85% for the topics of machine learning and 88% for the topic of data lifecycle. Similarly, all the materials and exercises used in the training were viewed as well suited, with the exercise on data lifecycle with Orange3 and AI-Bingo receiving slightly lower scores relatively to other exercises.

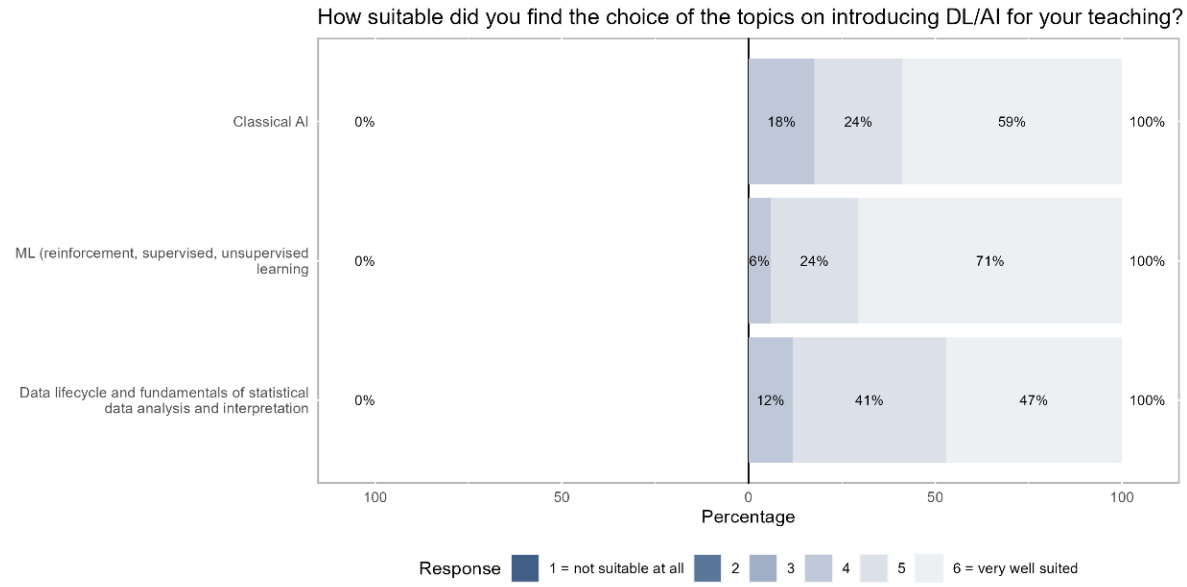


Figure 25 Distribution of the post results for the assessment of the training topics, 10.12.2022 Vilnius, post-evaluation survey, n=18

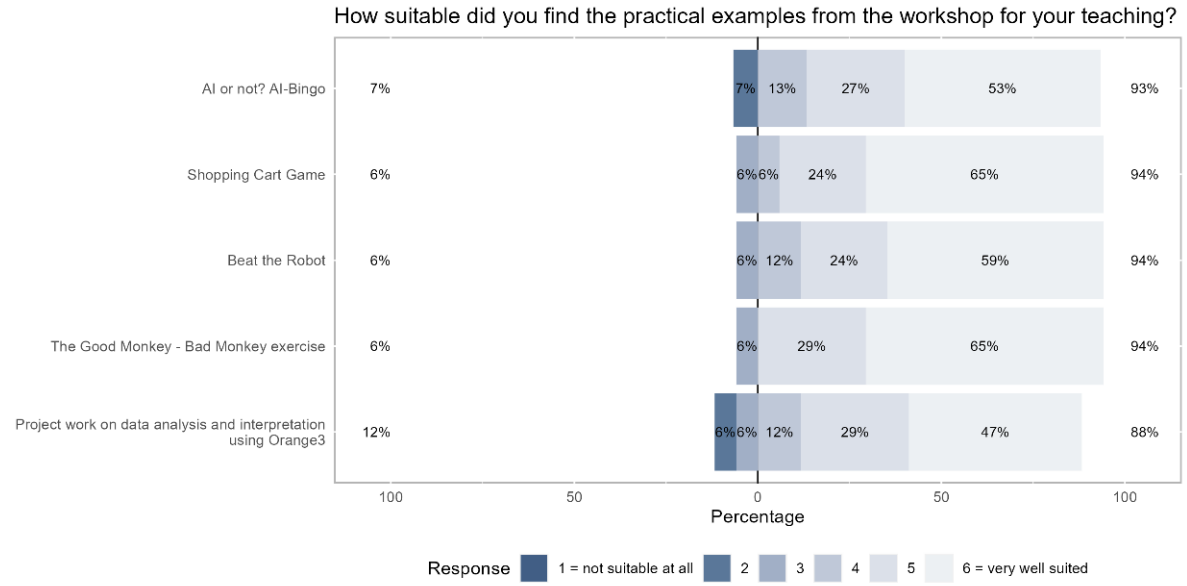


Figure 26 Distribution of the post results for the assessment of the training material/exercises, 10.12.2022 Vilnius, post-evaluation survey, n=18

Figure 27 shows additional items for teachers' post-training feedback, including training length and format. With respect to the length of the training, participants tend to agree that the training should have lasted longer. About 72% of the respondents selected the higher scores of "5" and "6". For the item "I wish more content would be covered in the training", 61% of the respondents lean towards strongly agreeing with this statement. There is, however, a strong consensus that the training did not

have too many teaching materials and that the interactive format was appropriate. 61% of the participants lean towards strongly agreeing with the statement, "The training showed me which competences I lack to teach the relevant content in class". The responses to the statement "I would need a lot of preparation to teach the topics and application examples in class" indicated a tendency towards agreement.

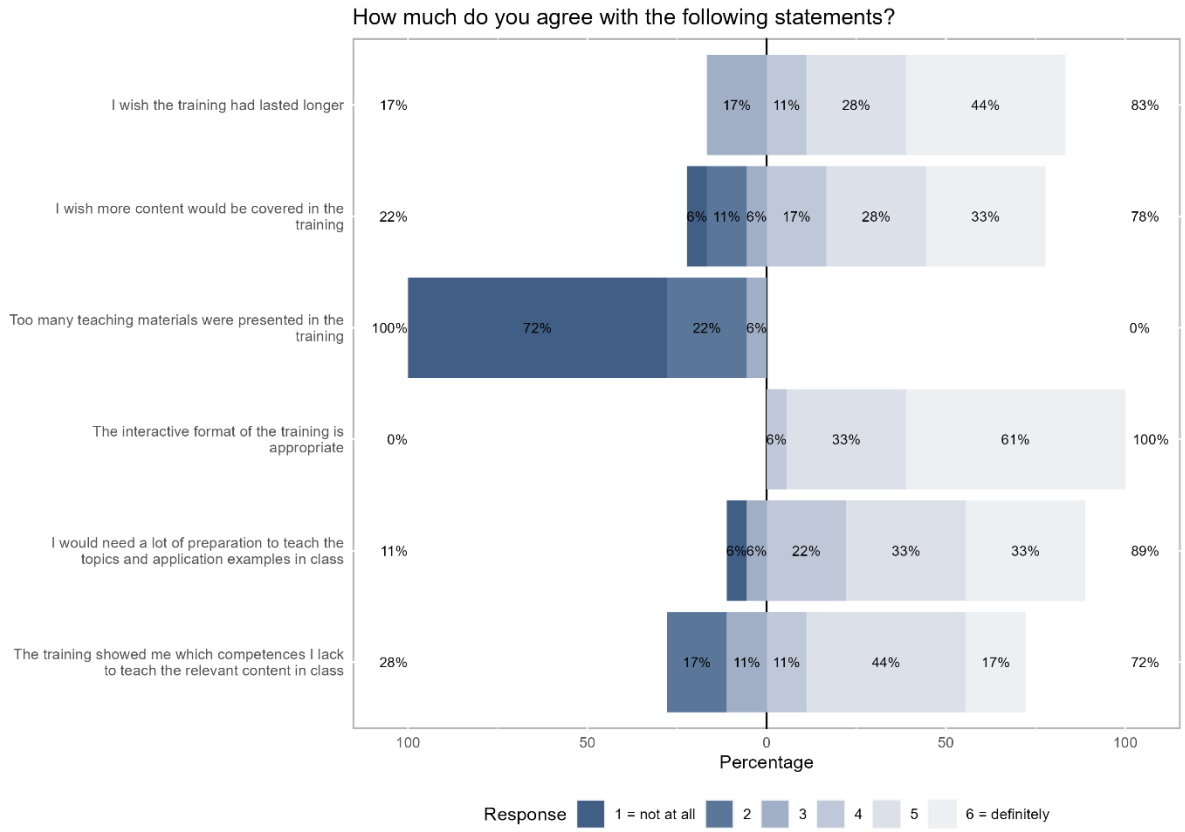


Figure 27 Distribution of the post results for the assessment of the training format and outcomes, 10.12.2022 Vilnius, post-evaluation survey, n=18

3.3 Qualitative results

3.3.1 Personal interviews

The themes listed below were deductively derived from the research questions and the structure of the interview guide. The annotations on the right refer to categories that were assigned inductively throughout our qualitative analysis process. Given the novel nature of the research questions, we chose not to limit our analysis to a purely deductive approach. Though our initial research questions and the interview guide shaped the general themes, it was the inductive analysis of participant responses that filled these themes with detailed and meaningful content.

A. Training:

Factors prompting training participation (prior to the training)

Regarding obtaining information about the training, the following was mentioned: assistance in finding courses through colleagues and receiving invitations through professional contacts.

Information was obtained through colleagues and professional networks

Expectations (prior to the training)

There was an anticipation of openness to all topics. The preference leans towards acquiring as much knowledge as possible. Participants wished to be updated about the present status of training subjects. Besides the need for information, there's also a desire for a shift in perspective. Additionally, participants wanted to learn more about Orange3, which they already worked with.

Building or updating knowledge

Teachers wanted to obtain new material, that can be used in the classroom.

Need for new material

Also, the focus was on the exchange with other training participants.

Exchange with colleagues was expected and desired

Design according to familiarity with DL and AI

The interviewees reported that the training was designed according to their previous experience and knowledge in the area of DL and AI.

Training well-aligned with pre-existing knowledge

There was no participant in the interviews who indicated a high familiarity with AI. Regarding high familiarity levels in DL and moderate/average familiarity in AI, some content was already known. However, both familiar and new content was explained very clearly: *"So, yes, I guess I liked the way that the concepts were presented and the way that the methods were shown, the activities that we can do with the students and probably the higher-level explanations when talking one-on-one. And, just to clarify not for the students, but just to know it myself."* (Participant C - 10.12.22, Vilnius) The content was also presented in a way that could be understood by both advanced and novice teachers.

Difficulty of the topics for the teachers themselves

Most answers are (1) differentiated, while (2) few answers clearly rank difficulty.

(1) It is classified depending on the subtopics and the further employment following the training: *"Sometimes it depends on the topic, I guess, because AI is a large, large topic that could be separated into different ones, and so is data literacy. [...] And, there are parts from like with the higher-level knowledge where I have to do some more research [...]."* (Participant C - 10.12.22, Vilnius)

Complexity levels are dependent on topics

(2) The difficulty arises not only from the magnitude of the topic of AI, but also from the complexity that goes along with it. This also makes it clear that there is much that is not yet known: *"For me, I know but I do not know enough about artificial intelligence."* (Participant F - 10.12.22, Vilnius). To deepen the knowledge of the training, further training is needed.

AI is viewed as complex and more training is needed

Orange3 was viewed as complex, as the time for this topic and tool in the training was too short.

Orange3 is viewed as complex and more training is needed

B. Teaching of DL and AI

Integration of DL and AI into teaching CS at the secondary level (prior to the training)

Integration can be divided into three areas: (1) lack of integration, (2) integration but not within the framework classes, and (3) integration in the classroom.

(1) Non-integration refers only AI content, and not DL. Reasons given for this are that one focuses on programming instead and secondly a lack of own knowledge on this topic.

No AI integration due to the lack of knowledge

(2) Integration of AI outside of framework classes refers to the inclusion of the AI topics into extracurricular projects: *"[...] because in our school, we used these projects for different experiments but it is like extracurricular activity more than lessons, so."* (Participant A - 10.12.22, Vilnius)

Integration of AI within extracurricular activities

(3) DL and AI were taught separately. In the case of DL, it also involves experimenting with data, e.g., within a research project*: "So, I*

Teaching DL and AI separately

challenge them [the students] to explore the data and find their own conclusions or raise the questions, raise the hypothesis, and then check them out using either really simple techniques like visualizing or we can also use some statistical formulas counting like correlation coefficients [...].” (Participant H - 10.12.22, Vilnius).

Integration of DL in math and CS classes

It can also be emphasized to use big data with many variables instead of small, artificial test data with only a few data points. In addition to basic work with data, topics such as data programs and internet security are also on the agenda. DL is integrated as part of statistics in mathematics, as a separate course, in CS via work with spreadsheets, and as *“part of information technology lessons [computer science lessons]”* (Participant F - 10.12.22, Vilnius).

Integration of AI in classroom

AI was included into the lessons via the integration of material such as the monkey-robot game and "Nim" for reinforcement learning. Furthermore, video documentaries were watched, and the ethics of AI were discussed. Basically, real-life references are used a few times, e.g. *“because kids use artificial intelligence every day in their lives”* (Participant D - 10.12.22, Vilnius). Machine learning was also used as a teaching topic, but not too often because the grade level there was relatively low.

Experience with integration of DL and AI into teaching CS at the secondary level (prior to the training)

Difficulty of explaining due to the lack of knowledge

Since the students belong to different class levels, there are sometimes difficulties in explaining. Furthermore, a connection to the lack of one’s knowledge is made: *“And yes, I could not really deliver, be a good teacher in this way and not understanding the topic myself.”* (Participant B - 10.12.22, Vilnius)

A further difficulty was seen in the lack of material, insofar as no 'official' teaching material and no textbooks were available, which made teaching more difficult.

Lack of material is a problem

In terms of DL, students understood how to interact with data in different ways: *"[...] what are the differences between data visualizations and different methods of collecting data, different sources of data, how to search for information on the internet and so on."* (Participant C - 10.12.22, Vilnius)

Students understood work with data

On the student side, it is described that the use of AI-unplugged exercises was well-received, and the students would likely be open to digital programs like ChatGPT in the future.

The analogue exercises worked well, no digital ones were used yet

Ability of (more) effective integration (after the training)

In higher grade levels, the training has only to a small extent enabled the ability to teach the topics: *"[...] you need to be a little bit more clever than your students. And you need to ask also/ you need to be prepared to answer the questions and to know how to learn and to learn others."* (Participant G - 10.12.22, Vilnius)

Training is not sufficient for integration for higher grade levels

Regardless of the grade level, the fact that the material is in English and not in Lithuanian is a hindrance to the use of the material. This makes it difficult not only to use the material, but also to evaluate it by the teacher.

Material being available only in English is a hindrance for the use

In other cases, the training is evaluated positively in terms of making the integration possible, but it cannot be implemented immediately, as materials require modification: *"And I think I will try to do*

Materials cannot be used 1:1

these. I will try to find something similar for my classes. So, in this way, I feel inspired." (Participant H - 10.12.22, Vilnius)

If integration is possible, it is possible at the lower grade levels.

Integration is possible for lower grade levels

Difficulty of imparting knowledge to school students (after the training)

The assessment of teachability ranges from (1) differentiated to (2) high.

(2) Content from the training is not considered to be suitable for older students, as they would quickly get bored with it. The needs and the level of the students must be considered, and the use of materials must be adjusted accordingly. Analogue exercises are not interesting for older students. However, the materials can be adapted. For example, Orange3 can be adapted for lower grades (4-5th) as a *"different way of implementing data use, like charts or just like to say how to get results in different form"* (Participant D - 10.12.22, Vilnius). Such modification of the material means a lot of work for teachers.

Materials must be tailored to specific grade levels

(3) Some teachers mention a high level of reachability without further commenting on it.

High level of teachability is mentioned

C. *Establishment and steps to integrate DL and AI:*

Framework curriculum integration

With a few exceptions, the feedback regarding framework curriculum integration is overwhelmingly positive. However, teachers mentioned the problem of the already overloaded curriculum, which makes it difficult to include new topics.

Positive attitude

The current framework curriculum is overloaded

Arguments for the inclusion into the framework

Inclusion into the framework curriculum was coalesced with various arguments. Regarding the topic of AI, it is necessary to take away fears and demystify the topic. Critical thinking can be promoted through the inclusion of the DL and AI topics. The knowledge of data interpretation should be increased. These topics can be linked to real-life scenarios, emphasizing their significance in how we perceive the world: *"It is a matter of thinking. It is not about, you know, just the programming itself or something. You know, just like how we go, how we see the world a little bit, you know?"* (Participant B - 10.12.22, Vilnius) Furthermore, as these subjects will become increasingly prevalent in society, it's imperative to integrate them now.

High number of teaching hours is needed

Not only are specific topics discussed, but also the time allocated to them in the classroom. There's a call for a significant number of instructional hours, beginning either in primary school or early secondary school. There's also a demand for more flexibility in the framework curriculum. Proper integration would streamline examinations, providing clarity to teachers regarding what can be assessed. A substantial number of teaching hours is essential.

Steps for anchoring in the classroom

Free and sufficient material should be available

Another step concerns access to sufficient material (websites, links, material, exercises such as from the TrainDL training). These should also be freely accessible to teachers. Ensuring the clarity on how content can be taught and with which materials is a basic requirement.

Teacher training should happen first

Another prerequisite is the appropriate training of: *"But beforehand, we have to know how we can use it, you know, it should be like two*

levels step teaching and also like how you approach, how you use it this moment, in the real life." (Participant B - 12.10.22, Vilnius)

Teacher education

Teacher training is essential

Teacher training is seen as an essential block for the new framework curriculum: *"So, I feel that these two things, a bit of teacher training and a bit of changes in curriculum will move the world."* (Participant A - 10.12.22, Vilnius). The need for continuous training is seen: *"First of all, train teachers, and the second is retrain teachers."* (Participant E - 10.12.22, Vilnius)

Some teachers see DL and AI as *"not integral"*, but *"should be covered at some capacity, maybe at the basic level"* (Participant C - 10.12.22, Vilnius). Another opinion views pedagogical competences central and sees them being neglected in the current teacher education so far.

School curriculum

Further training of teachers is central

To bring the topics into the school curriculum, the knowledge of the teachers is again brought into play in the form of further training. More materials and plans for teachers, also interdisciplinary ones (e.g., mathematics), are necessary.

DL and AI can be integrated in other subjects as well, not only in CS

Accordingly, one suggestion is to look for potential applications for DL and AI in the school curriculum in as many subjects as possible and to apply them there: *"[...] like in geography maybe talk about how the GPS works and how it could be, how it finds the best route or maybe in math where they use different solving methods or maybe even proof."* (Participant C - 10.12.22, Vilnius)

Integrating AI outside school curriculum

Since the school curriculum is based on the information technology curriculum (CS), it is necessary to introduce AI topics in the framework curriculum accordingly. Another way is to deliberately bring the topics outside the school curriculum: *"And I heard the others, like the projects outside of the curriculum that the schools are doing are really a good way for introducing these more complex topics to the students that are motivated and interested in them so that they can study more deeply into these topics."* (Participant H - 10.12.22, Vilnius)

Barriers

Teachers lack motivation to attend further trainings

(1) For the context of Lithuania, it is stated: *"But here in Lithuania, we do have this problem where teachers are not as enthusiastic about improving their competences and attending these trainings. [...] For me, yes, this is a really nice idea to improve my competences, but then I cannot really speak for the others here."* (Participant H - 10.12.22, Vilnius)

AI topics change quickly

Therefore, there is uncertainty about how to introduce these topics into the classroom and how to deal with them when there is no motivation to attend such trainings. This situation is reinforced by the quick development of the topics and the situation where students might have more knowledge than the teachers. On the teaching side, teachers might have fears and might feel intimidated concerning the AI integration.

Lack of training coalesced with strict school demands

Another obstacle is a lack of training/preparation on the part of the teaching staff. Furthermore, the absence of flexibility in planning is highlighted, especially when topics aren't included in the school syllabus, yet schools expect a strict adherence to the framework curriculum.

Limited time and high workload are major obstacles

One of the greatest obstacles is limited time and high workload: *"The biggest obstacle I see it is time because preparation for artificial intelligence sometimes takes huge amounts of time. You need to gather all data. You need to separate it. You need to understand what level your children are at, because you can gather a lot of data about artificial intelligence that they do not even have any ground to stand on with that information [...]."* (Participant D - 10.12.22, Vilnius) Without understanding from the teachers, there is no understanding from the students. Moreover, due to the lack of time, interesting 'side paths' cannot be taken when preparation for the exam sessions is such a high priority.

D. *(Possible) changes through the integration of DL and AI in the framework curricula:*

Students

Empowering students to make informed decisions and shape the future

Understanding new concepts and technologies, working with data leads to decision-making skills, which will also have an impact on later life outside school. The understanding also relates to everyday applications such as search engines (Google). Students are empowered to help shape the future.

Data security

The integration could lead to better data security decisions.

Social impact

It was suggested that students, as members of society, can promote transparency and have a broader societal impact through their interactions. *"[...] because our students are the part of society, they would talk more about transparency with parents, with their friends maybe a little bit, but have an impact."* (Participant E - 10.12.22, Vilnius)

DL and AI could be used to support students in their learning

Data can be used to provide support services, for example it can be used to predict patterns and, for example, to forecast whether students need help now or in the future. AI could personalise the

institutionalised teaching and learning process (lessons, tests, materials). It can also be used *"for the design of projects, or using it for learning languages, or anything like that"* (Participant H - 10.12.22, Vilnius).

School authorities, schools, and school curriculum

AI can be used at the school administrative level

Some teachers mentioned suggestions for the use of AI at the level of the school administration: *"[...] about how artificial intelligence can be used in the school administration, for administrative purposes, and what kind of data we collect [...] and how we can later make a decision and say find students with psychological difficulties or with learning problems and so on, not only using psychological models but also getting from the data we are collecting and other steps [...]."* (Participant B - 10.12.22, Vilnius) Linked to this is also the use of anticipating future conditions and influencing, for example, school interior design and curriculum, in short: *"Adapting the school making it more beneficial for the students and maybe the teachers."* (Participant C - 10.12.22, Vilnius)

Society

AI could be used to target social problems

AI problem-oriented solutions could be introduced, so that, for example, people with needs could be identified early and be provided with help.

Data literacy can help mitigating societal divisions

DL can contribute to solving and avoiding political problems: *"I feel that/ and also especially during the pandemic when there were discussions about the vaccines [...] society was really divided over the topic that probably could have been solved if data literacy and statistics/ understanding the data, interpreting the data, finding the information, evaluating whether these statistical inferences are correct or not, then this competence*

could have solved our even political problems here." (Participant H - 10.12.22, Vilnius)

The topics can lead to optimising workflows, making jobs and creative processes easier.

DL and AI could lead to more efficiency and creativity at work

E. Training feedback and potential for improvement:

Length

All teachers found the duration acceptable. Some proposed extending the training into a series spread over several weeks or a month, allowing for more in-depth knowledge acquisition and the opportunity to test the training content with students intermittently.

7 hours is a good length, but a series of trainings could be tried out

Content

Topics (Classical AI) and individual exercises ("Good Monkey - Bad Monkey" Game, "Shopping Cart" Game) were praised. The content was rated as useful.

Content was positively praised

Orange3 was generally viewed positively and found interesting. However, some feedback included desires for a more in-depth exploration of the program, comments that the allocated time was too brief or that the content was presented too hastily. There were also suggestions to pair Orange3 with other software and to explore more features beyond just decision trees.

Orange3 requires more time

One teacher emphasized the importance of data preparation, remarking: *"And I think it is, I would say that at least people, they must analyse this data but probably to clean themselves this data or to prepare themselves this data. Otherwise, it is not so easy to understand how we are getting to this data for this Orange tool example."* (Participant A - 10.12.22,

Cleaning and evaluating data can be useful activity

Vilnius) The preparation can then take place in a Google form, for example. It would also be helpful to know how to evaluate data (good/bad data).

Apart from this, other topics of interest include neural networks, real-time face and voice recognition in the training itself, e.g., for gender assignment.

Other AI topics of interest: neutral networks, real time recognition of faces and voices

There's a need for a more detailed understanding of AI functioning (like in video games) to effectively convey this to students. Additionally, exploring how AI can be integrated across diverse subjects such as art, music, and literacy is essential, offering multiple perspectives on its application.

Deeper understanding of AI is needed to convey knowledge to students

Format

The training format was highly appreciated, seen as a well-structured extended lecture complemented with exercises, and was perfectly aligned to the target audience's needs: *"For today it was very good because it is clear that the lecturer did know what her audience was."* (Participant F - 10.12.22, Vilnius)

Format received positive feedback

Interaction

The interaction was highly valued A good balance was achieved with the lecture component, ensuring that questions were addressed and tools were explained in an accessible manner for all.

The level of interaction was praised

TrainDL training (other factors, overall view)

The presentation's clarity stood out: *"But like it was presented today, everyone liked it. I think that it was a good form and seminars like this are*

Presentation of material was clear

acceptable by teachers, they will understand everything because everything is made step by step. You do not just dive into the course without any preparation. So, I like it very much, and I think it is a good teacher training variation." (Participant D - 10.12.22, Vilnius)

The structure of the training was good

The training provided a beneficial introduction to the topics and successfully introduced new content to participants. The overall structure of the training was praised, emphasizing its student-centric approach: : "[...] from the kids, like for the monkeys, that was really truly interesting and there was like understandable for us as young people for the person." (Participant G - 10.12.22, Vilnius)

More small group activities could be beneficial

It was suggested to work more in small groups: *"Maybe smaller groups could help sometimes. Like going into smaller groups of like five people or something like that, some maybe workshops of methods and so on."* (Participant C - 10.12.22, Vilnius)

Implementation plans with varying difficulty levels are needed

Regarding integration, there's a desire for concrete lesson implementation plans with varying difficulty levels (*"simple, normal, and higher"*) (Participant E - 10.12.22, Vilnius).

Material in Lithuanian is needed

Material in Lithuanian is requested, as the translation of the material into English is time-consuming.

Other trainings (not TrainDL)

More teacher training is needed

Ideally, more events similar to TrainDL are encouraged, not just for the targeted teacher group. For CS teachers, an expansion into AI-specific subjects is recommended: *"I guess, yes, teacher training maybe some, not mandatory, but maybe optional courses that would be concentrated on methods and on maybe two different subject teachers.*

Where there is AI for like geography teachers and then maybe a separate course where there is AI in, I do not know, maybe in technology and so on."
(Participant C - 10.12.22, Vilnius)

DL can be taught outside of CS

Deep learning and neural networks, for example, could then be dealt with in a topic-related manner. For DL, it is suggested to make the applicability in real and school life clear as well as to design further training for other subjects apart from CS, as they also work with data.

3.4 Summary

The participants in the CS training in Vilnius were predominantly in-service teachers. Their second subjects were diverse, with mathematics being the most common. Many of the participants had at least some experience of teaching DL in their classes. The training had a high representation of women, which aligns with the high ratio of female teachers in Lithuania in general.

The findings from the pre- and post-policy evaluation survey indicate that the training had a positive impact on teachers' perceived competences on how to use AI in class. The Wilcoxon signed-rank test did not indicate a statistically significant difference between the pre- and post-training for the item on how to use DL in class, albeit the average pre-training level was already relatively high. The willingness to invest time and effort to incorporate DL and AI was generally high. Participants expressed a tendency towards agreement with the statement that they have acquired sufficient competences to teach the learned content in class.

Self-report knowledge questions showed that participants improved their understanding of both DL and AI concepts. However, objective knowledge test confirmed that only for the AI part. The distributions of the pre- and post-scores were quite

similar and the Wilcoxon signed-rank test showed no statistically significant difference for the DL knowledge test scores.

The post-evaluation survey provided further insights on teachers' reactions towards the materials and exercises of the training. All the materials and exercises used in the training were viewed as well suited, with the data lifecycle exercise using Orange3 and AI-Bingo receiving slightly lower scores relatively to other exercises. With respect to the length of the training, participants tend to agree that the training should have lasted longer. There is, however, a strong consensus that the training did not have too many teaching materials and that the interactive format was appropriate. Participants, on average, indicated agreement with the statement that the training successfully highlighted the competences they need to enhance to teach the relevant content in class. Furthermore, the responses also demonstrated a tendency towards agreement with the statement that substantial preparation would be required to effectively deliver the topics and application examples in class.

All teachers acknowledged high importance DL and AI competences for both students and society. Participants express a tendency to believe that AI and DL content is lacking in the current CS framework curriculum. This perception does not change immediately after the training, as there is no statistically significant difference between the pre- and post-results. Similarly, there is no significant difference for the pre- and post-results for the items: "In the future, teaching DL/AI will provide added value to students." Additionally, the item on the societal importance of both topics collected only in the post-survey, revealed a very high level of agreement that DL and AI have enough societal importance to be incorporated into the curriculum. Findings also suggest that teachers have positive expectations about students' interest and their ability to generate student excitement and engagement for DL and AI topics.

Overall, the quantitative findings (albeit limited by a small sample size) suggest that the 7-hour CS training in Vilnius enhanced teachers' perceived competences on how to use AI in class, as well as their understanding of AI-concepts. The DL-content

of the training was more challenging. The findings also suggest that while participants report positive attitudes towards DL and AI, we cannot conclude that the training had an impact on that. The high level of agreement among participants regarding the training's ability to reveal the competences teachers lack, along with the shared perspective that a longer duration would be desirable, suggests limitations of a single 7-hour training for the comprehensive integration of DL and AI topics into teaching.

The qualitative findings showed that the training was, for the most part, designed in accordance with the participants' prior experiences, albeit teachers perceived the topics as complex and wish more training. The time allocated to Orange3 and data lifecycle was viewed as too short. They also state, that materials must be tailored to specific grade levels and be provided in Lithuanian language. Some of the topics had been partially integrated into the classroom previously, but within this integration AI received less focus, due to the lack of teacher proficiency in that topic.

Opinions regarding the feasibility of integration after the training are mixed. For some interviewees, the training alone doesn't sufficiently prepare them for immediate integration. For others, immediate integration is possible, but this primarily applies to basic concepts and lower grade levels.

The integration of the framework curriculum for DL and AI is viewed positively and should be supported by an adequate number of instructional hours. The goal should be to weave these subjects into a wide range of disciplines, not limited to CS. However, this requires teachers to have the necessary knowledge and access to relevant teaching materials. There's generally consensus on the need for corresponding teacher education. Barriers to introducing DL and AI into the classroom include the rapid pace of these topics, insufficient time and proficiency on the part of the teachers, and the already overloaded framework curriculum.

Introducing DL and AI into the framework curriculum is seen overwhelmingly positively, underscoring their high relevance and importance. However, the challenge lies

in scaling up teacher training, as many educators exhibit limited motivation for attending extended training sessions. It's imperative to devise strategies that incentivize teachers to engage in comprehensive training, transitioning beyond mere one-off sessions to a series that allows them to apply their learnings in a classroom setting in real-time.

4. CS in-service training, Austria

4.1 Sample

The sample of the pre-survey included 19 participants from Vienna (11 participants), Lower Austria (5 participants), Upper Austria (1 participant), Styria (1 participant), and Voralberg (1 participant). Figure 28 describes socio-demographic data of the participants as reported in the pre-survey. The average age of the participants in the training was 41 years old, ranging from 26 to a maximum of 61 years. Three out of 19 participants were women (16%). One person has chosen a non-binary category, "diverse". All the participants were fully trained in-service teachers. Everyone with one exception indicated that they teach at the lower secondary and upper secondary level. One participant selected the category "other". Out of 19 participants, nine were teaching at a general education secondary school. Others reported types of schools included: vocational higher education institution, secondary school, and polytechnic school.

In Austria, informatics is initially taught as part of other subjects at lower secondary level and later introduced as a separate subject (European Commission / EACEA / Eurydice). Based on the 2020 indicators released by the OECD, the share of women for all subjects at the lower and upper secondary level is similar to that of Germany - 65% (OECD 2023c). In terms of age group distribution for secondary school teachers, 14% of teachers are under 30. The 30-39 age bracket accounted for 19% of the workforce, while the 40-49 and 50-59 age groups represented 23% and 34% respectively. 12% of teachers are 60 years old or older (OECD 2023c). Unfortunately, data subdivided

by age and teaching subjects and specifically for CS teachers were readily accessible at the time of this report.

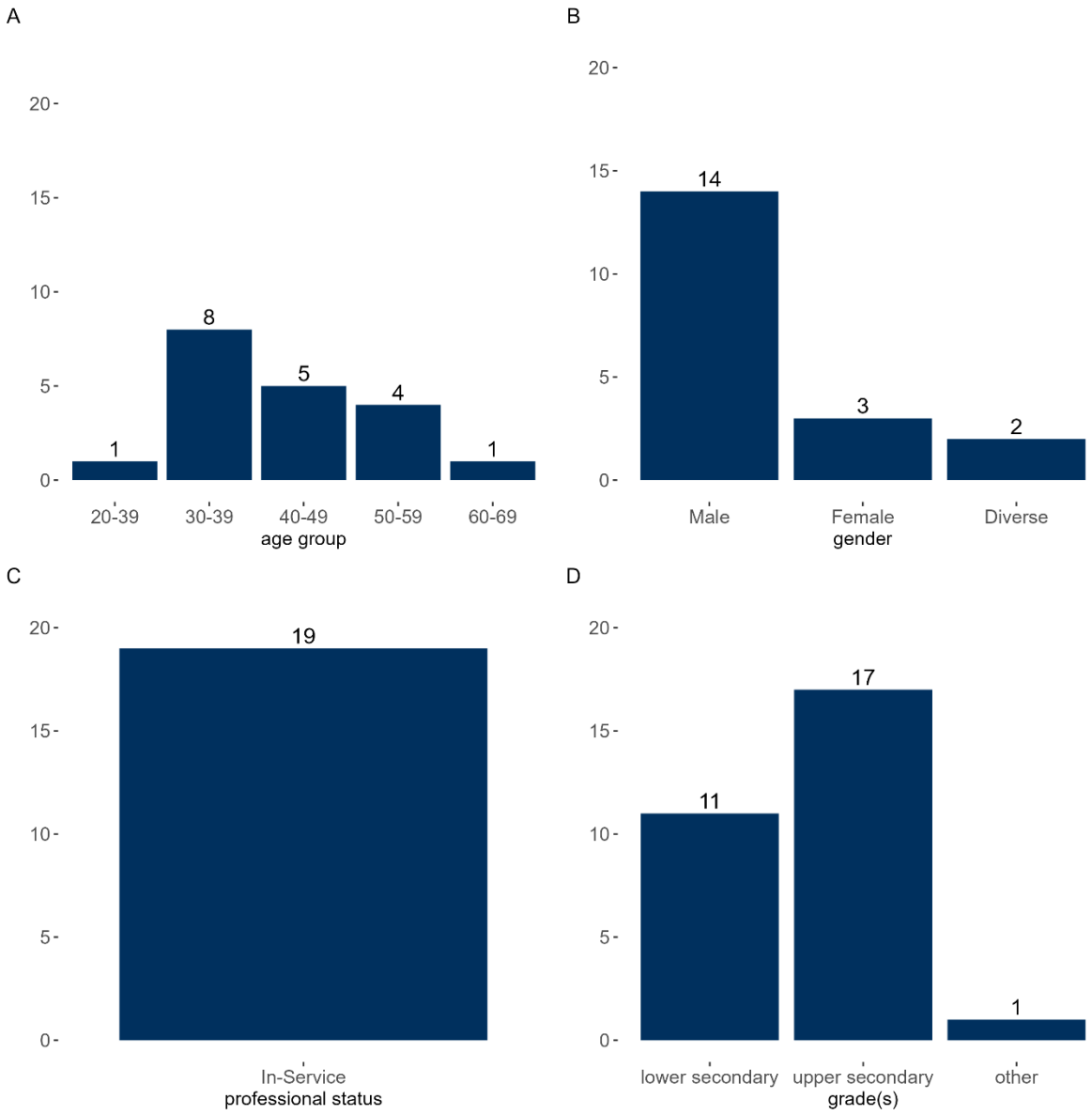


Figure 28 Socio-demographic data of the participants, 31.01.2023 Vienna, pre-evaluation survey, n=19

Except for one person who was teaching ethics and religion (albeit in the context of CS), all the participants were teaching CS as a subject, with the following additional/second subjects: digital basic education, mathematics, ECDL (European Computer Driving Licence), physics, chemistry, economics, ethics, geography, geometry, music, psychology and philosophy, and other. According to the pre-survey, digital

basic education and mathematics were the most common second subjects among the participants. Prior to the workshop, most participants had some experience teaching DL in class. As Figure 29 shows, while 12 of the participants never taught AI in class, only 5 out of 19 respondents reported never conveying knowledge about DL in class.

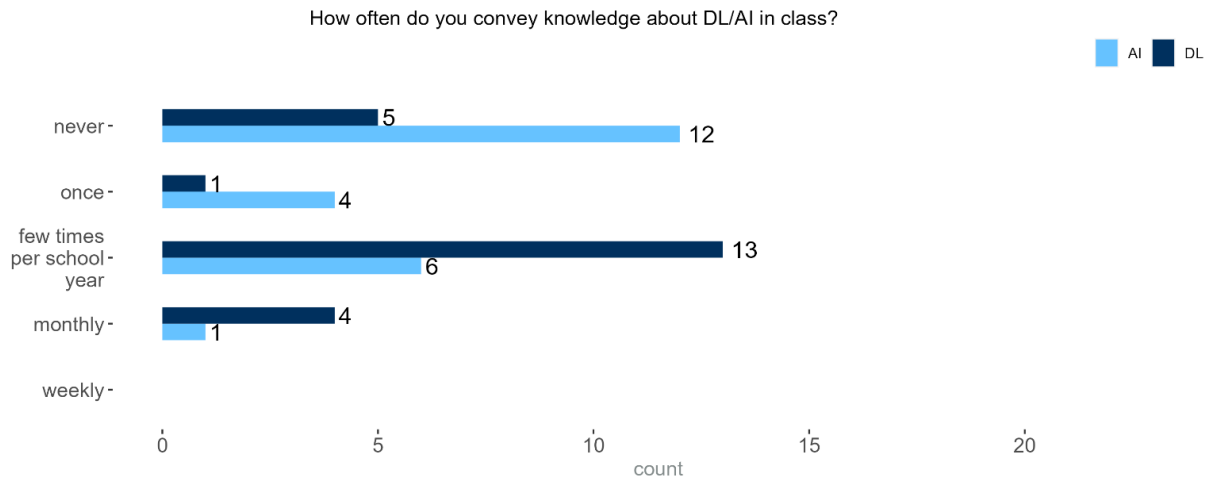


Figure 29 Experience with DL/AI, 31.01.2023 Vienna, pre-evaluation survey, n=19

4.2 Quantitative results

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

In both the pre- and post-surveys, participants were asked to rate their perceived competences in how to use DL and AI content in the classroom using a 6-point scale. Figure 30 and Table 13 demonstrate that on average respondents showed an increase in perceived competences for both DL and AI.

For the DL part, the median score increased from 2.5 to 4.0. The Wilcoxon signed-rank test showed a statistically significant difference between the pre- and post-distributions at the 5% level. Additionally, the interquartile range (IQR) notably decreased, indicating a reduction in the variability of post-training responses. For the AI part, the median score increased from an initial score of 3.5 to a post-training score of 4.5. The Wilcoxon signed-rank test indicated a statistically significant difference

between the pre- and post-distributions at the 1% level. The spread of the central 50% of the data also became smaller.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
I know how to use content about DL in the classroom	2.5	4.0	1.0	4.0	4.0	4.0	1	3	5	5
I know how to use content about AI in the classroom	3.5	4.5	2..25	4.0	4.0	5.0	1	3	5	5
I am willing to invest time and effort to incorporate DL into my teaching	-	5.0	-	3.75	-	5.25	-	3	-	6
I am willing to invest time and effort to incorporate AI into my teaching	-	5.0	-	4.75	-	6.0	-	4	-	6
After the training I have gathered enough competences to teach the learned content in class	-	4.0	-	3.0	-	4.25	-	2		5

Table 14 Summary statistics of pre- and post-results for the survey items on teachers' perceived competences to use DL and AI content in class and post-survey results for additional items, 31.01.2023 Vienna, pre-and post-evaluation survey, n=14

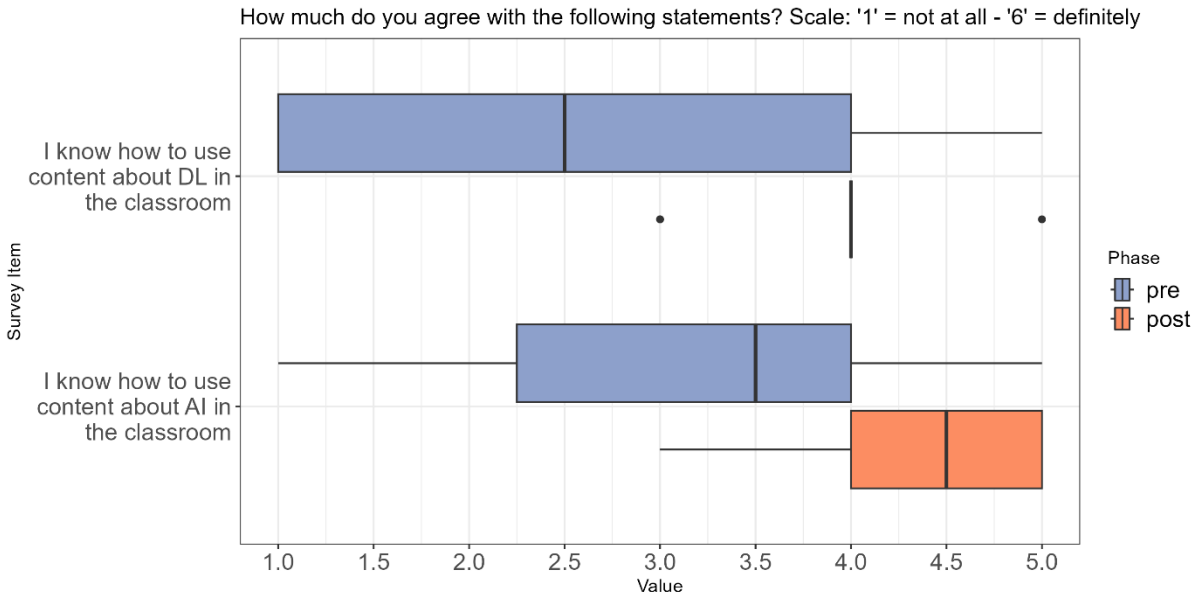


Figure 30 Boxplot comparison of pre- and post-results for the survey items on teachers' perceived competences to use DL and AI content in class, 31.01.2023 Vienna, pre- and post-evaluation survey, n=14

Figure 31 and Table 14 provide an overview of additional survey items that were measured only in the post-evaluation of the survey. Participants were asked about their agreement with the statements: "I am willing to invest time and effort to incorporate DL/AI into my teaching". For both DL and AI, most participants selected categories "5" and "6", indicating a strong willingness to invest time and effort to incorporate DL and AI into their teaching. After the training, participants were asked about their agreement with the statement: "After the training, I have gathered enough competences to teach the learned content in class". The statistics for this post-survey item suggest a moderate level of agreement among respondents.

Overall, the findings indicate that the CS training enhanced teachers' perceived level of competences on how to use both DL and AI in their teaching. In addition, the post-survey measures revealed a positive inclination among participants to dedicate time and effort towards integrating DL and AI into their teaching. On average, participants expressed a moderate level of agreement with the statement that they have acquired enough competences to teach the learned content in class after the training.

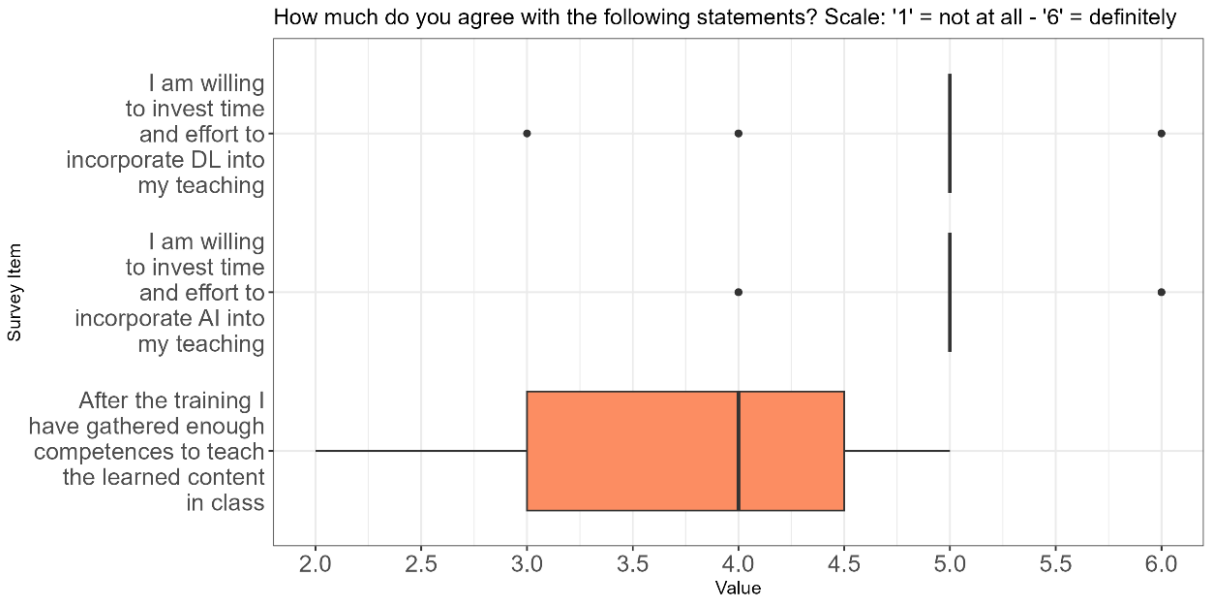


Figure 31 Boxplot of post results for the survey items on teachers' perceived competences and willingness to invest time and effort to incorporate DL and AI into their teaching, 31.01.2023 Vienna, post-evaluation survey, n=15

4.2.2 Understanding of DL/AI topics presented in the training

For both DL and AI self-assessment questions, there is an increase in median scores, with differences between the pre- and post-scores being statistically significant only for the AI set of questions at 1% (see Figure 32 and Table 15). The IQR reduced for the post-training values for both DL and AI values.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Self-assessment for DL	2.5	3.0	2.0	3.0	3.0	3.0	1	3	4	5
Self-assessment for AI	2.0	4.0	2.0	4.0	3.75	4.0	1	3	4	5
Knowledge test DL	1.7	2.2	0.6	1.5	2.15	2.5	0.0	3.0	1.0	3.2
Knowledge test AI	4.5	6.5	2.25	5.5	5.75	7.0	0.0	8.5	3.0	9.5

Table 15 Summary statistics of pre- and post-results for self-assessment and knowledge test, 31.01.2023 Vienna, pre- and post-self-assessment and knowledge test, n=11

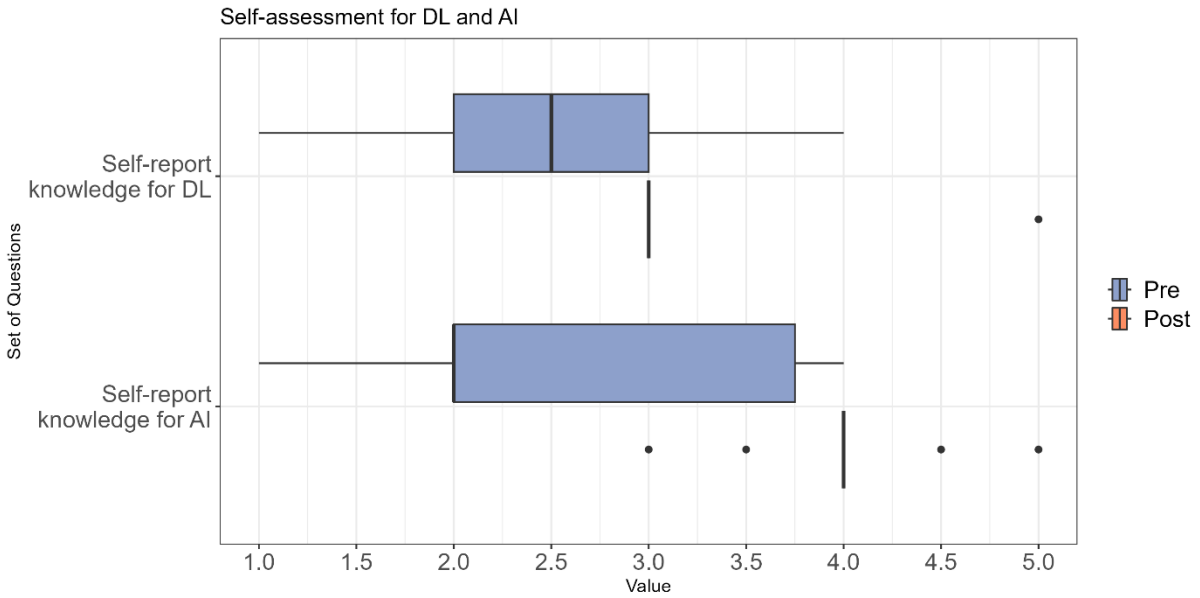


Figure 32 Boxplot comparison of pre- and post-results for self-report knowledge questions on DL (15 questions) and AI (10 questions), 31.01.2023 Vienna, pre- and post-self-assessment and knowledge test, n=11

Immediately after the workshop, participants demonstrated a higher median score and smaller IQR for the AI objective knowledge questions compared to the pre-training results (see Table 15 and Figure 33). The Wilcoxon signed-rank test indicated a statistically significant difference between the pre- and post-distributions at the 1% level. The post-training increase in the understanding of the DL concepts as measured via objective questions was not as notable. Yet, the median score increased slightly, IQR reduced, and the Wilcoxon signed-rank test indicated a statistically significant difference between the pre- and post-distributions at the 5% level. Given that the maximum score for DL is 4.0 and for AI 10.0, on average respondents had difficulties grasping all the concepts introduced in the training.

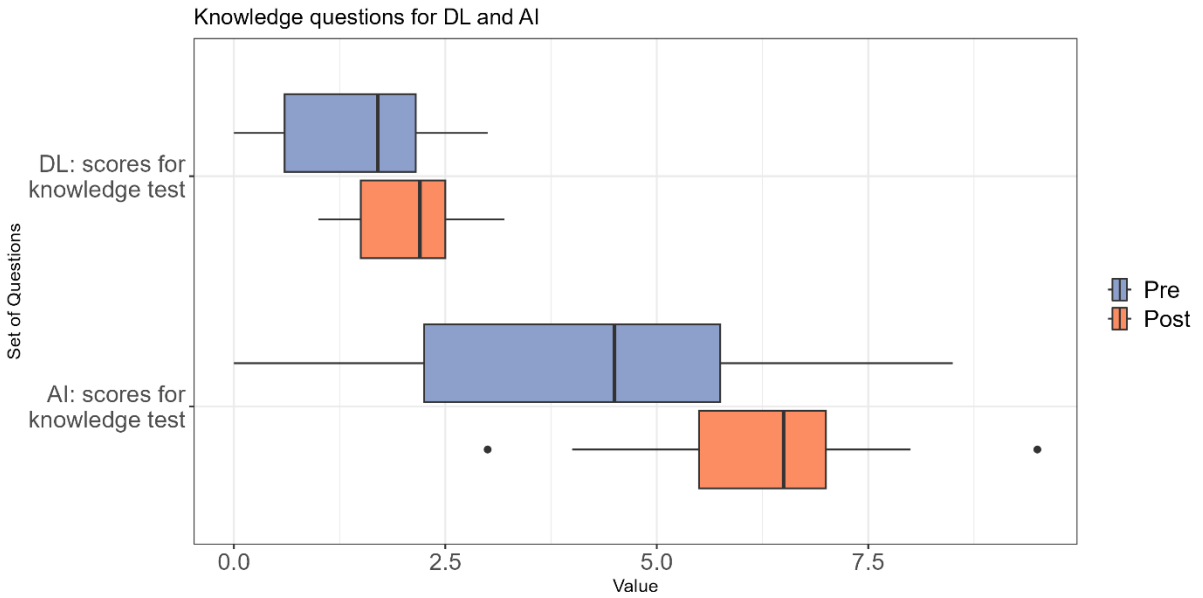


Figure 33 Boxplot comparison of pre- and post-results for knowledge questions on DL (4 questions) and AI (10 questions), 31.01.2023 Vienna, pre- and post-self-assessment and knowledge test, n=11

4.2.3 Attitudes towards DL/AI

Table 16 and Figure 34 describe the distribution for the survey items on teachers' perception of importance of DL and AI, which were measured via both pre- and post-policy experimentation surveys: "In the future, teaching DL/AI will provide added value to students" and "I think the content of DL/AI is missing in the current framework curriculum of CS."

For both DL and AI items, participants express a tendency to believe that either content is lacking in the current CS framework curriculum, although there is no strong consensus among the responses. There is no statistically significant difference between the pre- and post-results, albeit the post-training values for the AI curriculum item reduced from 5.0 to 4.0 with the spread of the values being quite high.

Similarly, there is no statistically significant difference for the pre- and post-results of the items: "In the future, teaching DL/AI will provide added value to students." However, participants indicated a notably high score for both items in the pre-survey

as well as the post-survey, indicating a pre-existing high level of perceived importance of DL and AI for teaching prior to the training.

Question	Median		Q1		Q3		Min.		Max.	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
In the future, teaching DL will provide added value to students	5.0	5.0	5.0	5.0	6.0	6.0	4	2	6	6
In the future, teaching AI will provide added value to students	6.0	5.0	5.0	5.0	6.0	5.0	4	3	6	6
I think the content of DL is missing in the current framework curriculum of computer science	4.0	4.0	3.0	3.0	5.0	5.0	2	2	6	6
I think the content of AI is missing in the current framework curriculum of computer science	5.0	4.0	4.0	3.0	6.0	5.0	2	2	6	6
The topic of DL is of enough societal importance to integrate it into the curriculum	-	5.0	-	5.0	-	6.0	-	4	-	6
The topic of AI is of enough societal importance to integrate it into the curriculum	-	5.0	-	5.0	-	6.0	-	5	-	6

Table 16 Summary statistics of pre- and post-results for a set of survey items on teachers' perception of importance of DL and AI, 31.01.2023 Vienna, pre- and post-evaluation survey, n=13

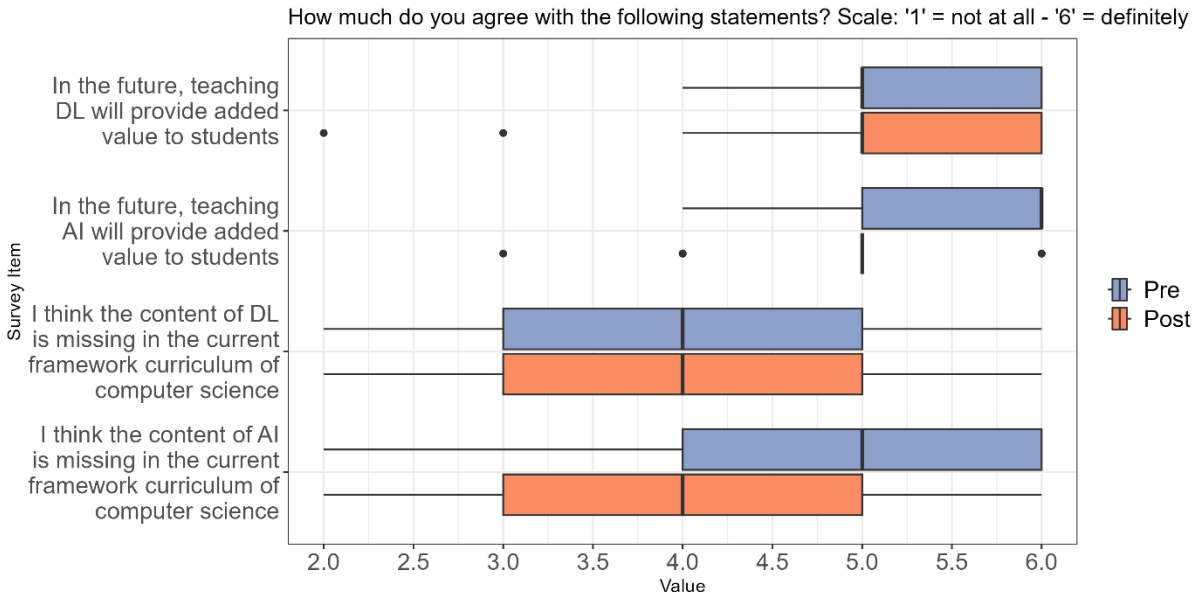


Figure 34 Boxplot comparison of pre- and post-results for a survey item on teachers' perception of importance of DL and AI, 31.01.2023 Vienna, pre- and post-evaluation survey, n=13

Similar to the trainings in Berlin and Vilnius, an additional survey item administered only in the post-survey revealed a strong consensus regarding the societal significance of DL and AI, with participants highly agreeing that these topics are of enough societal importance to be integrated into the curriculum (see Figure 35 and Table 16).

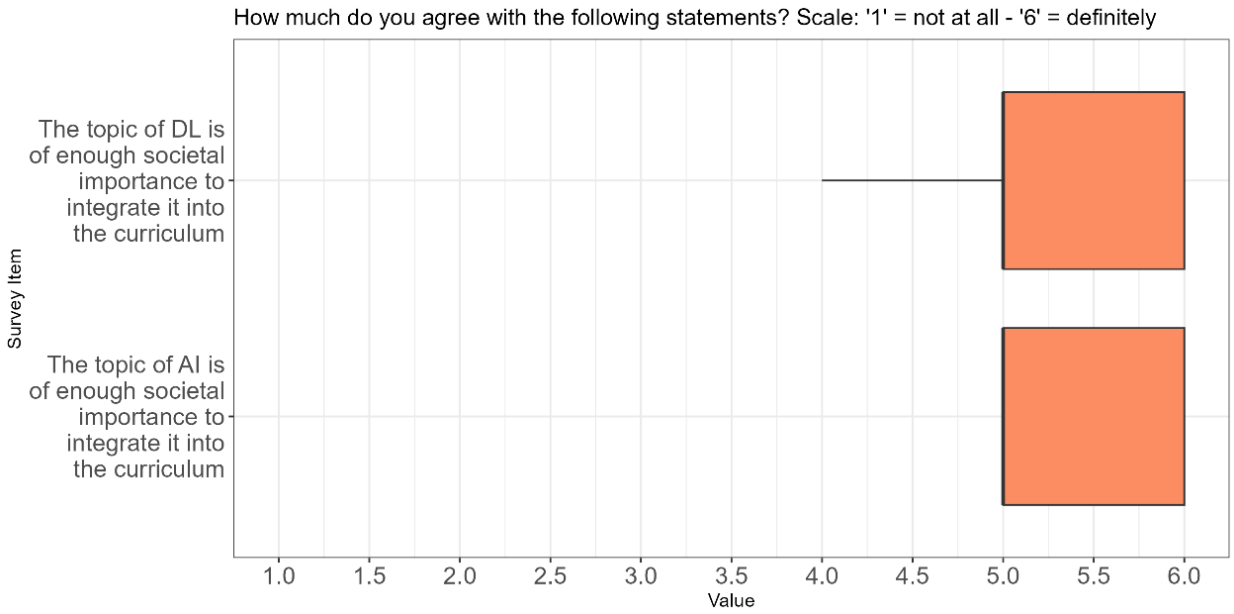


Figure 35 Boxplot comparison of post results for the survey items on teachers' perception of societal importance of DL and AI for integration into the curriculum, 31.01.2023 Vienna, post-evaluation survey, n=15

4.2.4 Perceptions of student engagement

In the post-survey, teachers were posed targeted questions to evaluate their anticipations regarding students' interest and their own perceived capacity to foster student engagement and enthusiasm in DL and AI subjects. These questions included statements such as "I can imagine my students show enthusiasm for the overall topic of DL/AI" and "I am confident that I can get students excited about DL/AI-projects."

The finding presented in Figure 36 and Table 17 showcase a generally optimistic anticipation of students' interest for both DL and AI. Figure 37 and Table 17 also indicate a notably high level of confidence among teachers regarding their ability to motivate students for DL and AI projects.

Question	Median	Q1	Q3	Min.	Max.	Scale
I can imagine my students show enthusiasm for the overall topic of DL	5.0	4.0	5.0	3	6	1-6
I can imagine my students show enthusiasm for the overall topic of AI	5.0	5.0	5.0	3	6	1-6
I am confident that I can get students excited about DL-projects	3.0	3.0	3.0	3	4	1-4
I am confident that I can get students excited about AI-projects	3.0	3.0	3.0	3	4	1-4

Table 17 Summary statistics of post results for the survey item "I can imagine that my students will show enthusiasm for the overall topic of DL/AI", 31.01.2023 Vienna, post-evaluation survey, n=15

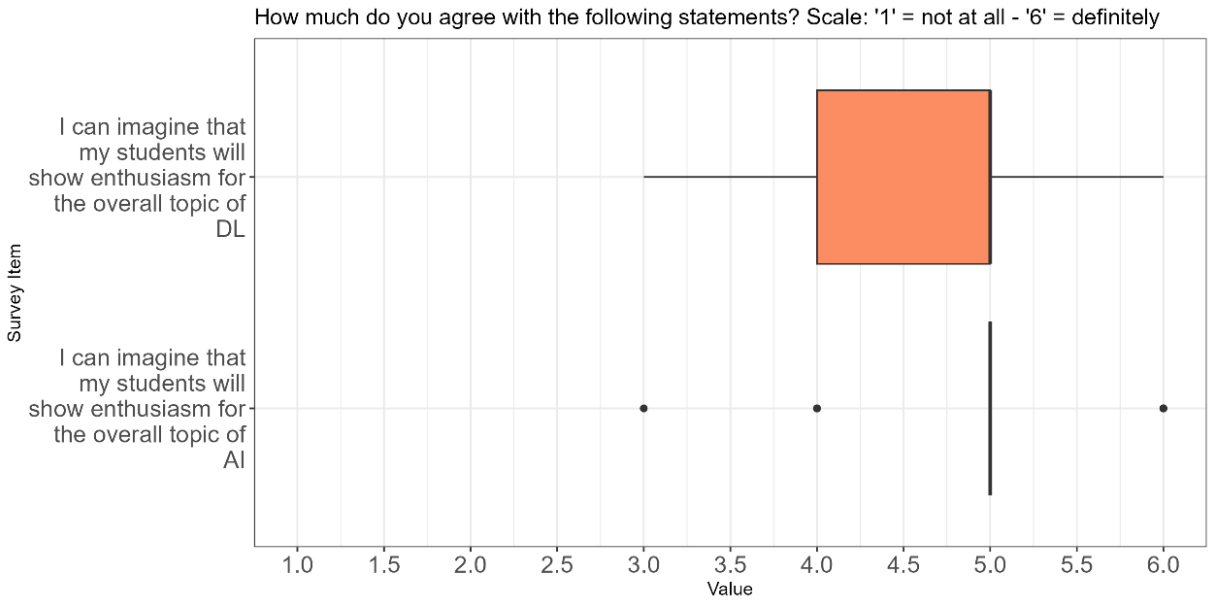


Figure 36 Boxplot of post results for the survey item on perception of student engagement, 31.01.2023 Vienna, post-evaluation survey, n=15

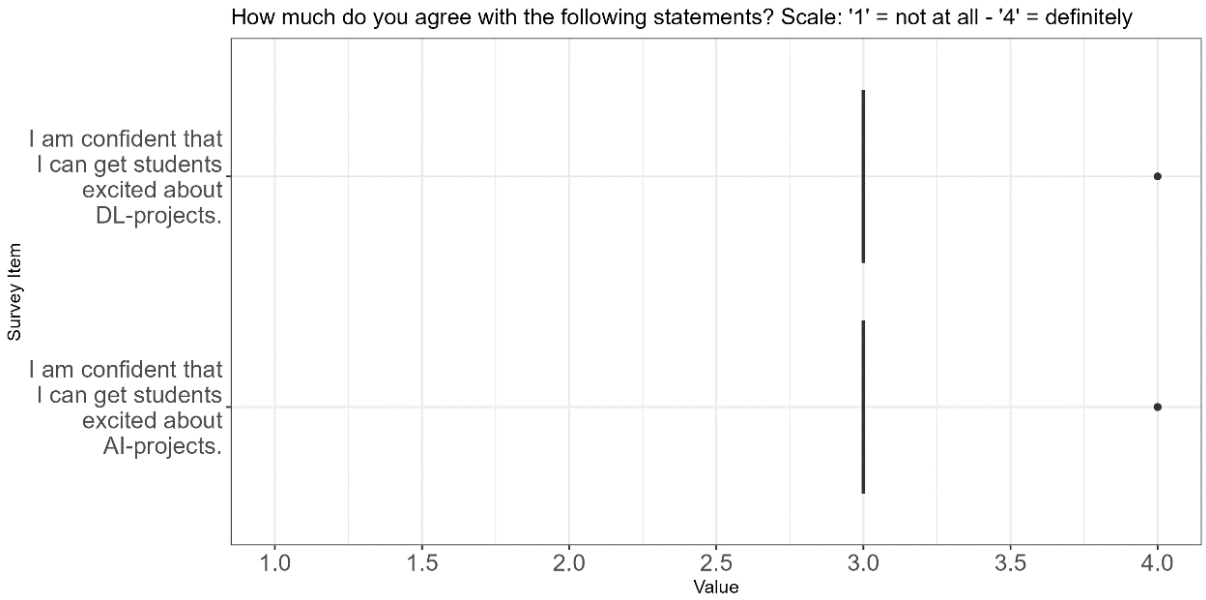


Figure 37 Boxplot of post results for the survey item "I am confident that I can get students excited about DL/AI-projects", 31.01.2023 Vienna, post-evaluation survey, n=15

4.2.5 Feedback on the learned content and format

The post-survey included a series of questions to assess participants' reactions to the topics, materials and exercises covered in the training. Figure 38 and Figure 39 present the feedback. The findings indicate that most participants consider the presented topics of classical AI and machine learning (ML) to be rather suitable for their teaching. Approximately 87% of respondents selected a score of "4", "5" or "6" (where "6" indicates a high level of suitability). However, the topic of data lifecycle was viewed as less suitable, with only 60% of respondents selecting a score from "4" to "6." Similarly, when evaluating the materials and exercises used in the training, it was observed that the exercise on data lifecycle using the Orange3 tool did not yield as positive results as the other exercises. While at least 70% of respondents rated all other exercises as "4", "5" or "6" indicating high suitability for teaching, only 53% of participants gave the data literacy exercise with Orange3 the same rating. The results suggest that participants highly appreciate and perceive classical AI and ML topics as suitable for teaching, whereas there is a need for improving the delivery of the data lifecycle topic and the DL exercise with Orange3.

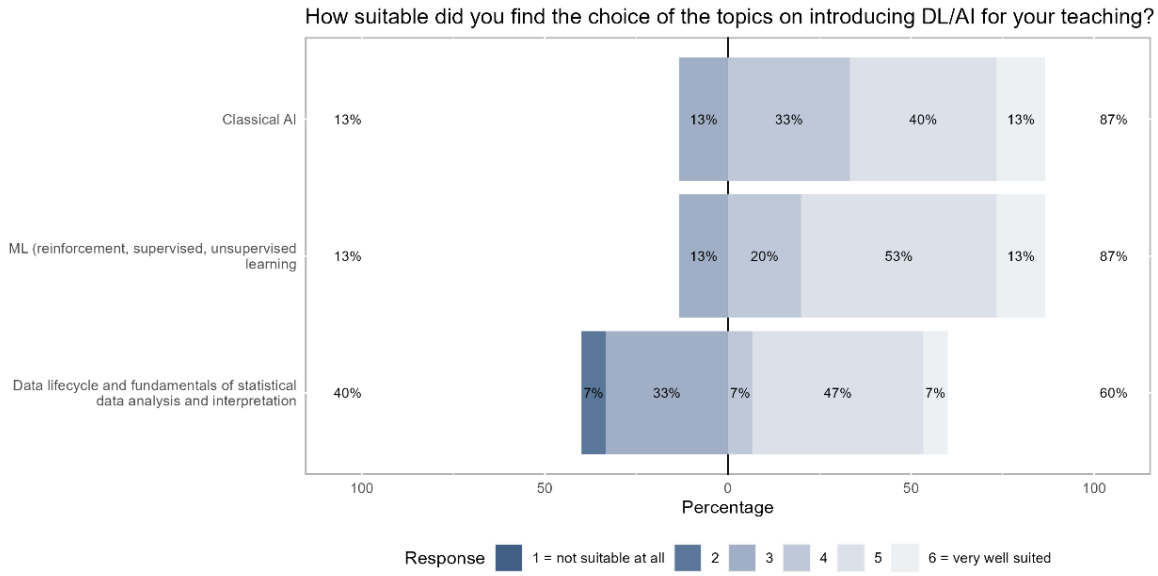


Figure 38 Distribution of the post results for the assessment of the training topics, 31.01.2023 Vienna, post-evaluation survey, n=15

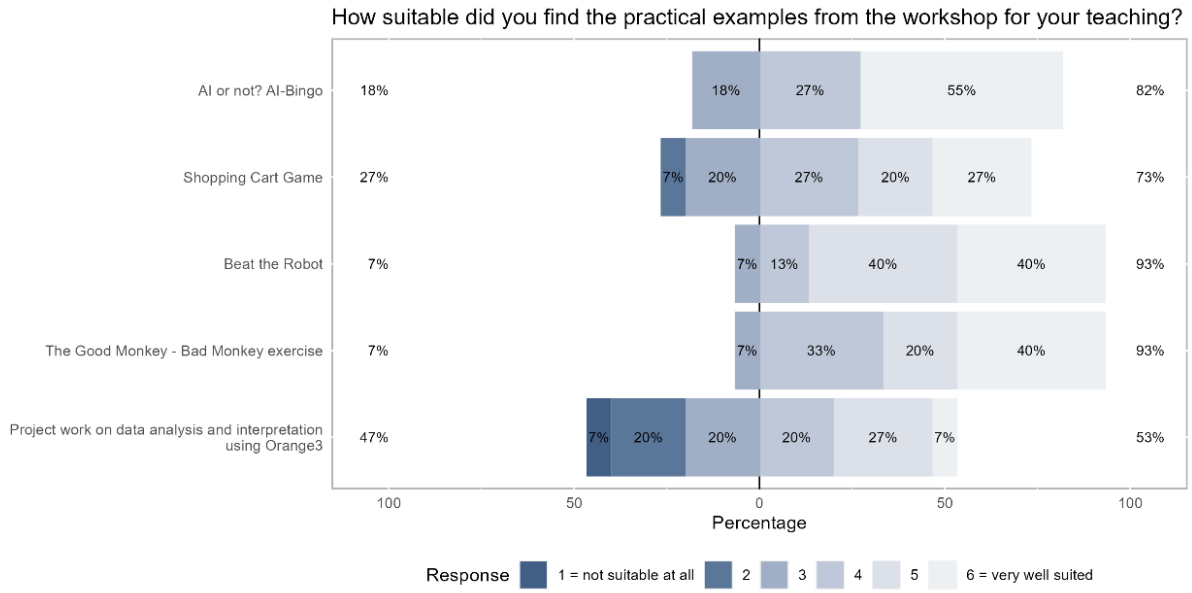


Figure 39 Distribution of the post results for the assessment of the training material/exercises, 31.01.2023 Vienna, post-evaluation survey, n=15

Figure 40 shows the results of teachers' feedback on training format and length. With respect to the length of the training, participants are somewhat divided in their opinions, and there is no indication of strong consensus for the statement that the training should have lasted longer. Of the 15 participants, five selected the higher

scores of "5" and "6", whereas three respondents opted for the lower scores of "1" and "2". The remaining participants selected scores within the intermediate range. The results regarding the item "I wish more content would be covered in the training" are also rather mixed, with eight respondents strongly leaning towards disagreeing with this statement and three respondents strongly agreeing with it. There is, however, a strong consensus that the training did not have too many teaching materials and that the interactive format was appropriate. Most of the participants lean towards agreeing with the statement, "The training showed me which competences I lack to teach the relevant content in class". The statement "I would need a lot of preparation to teach the topics and application examples in class" had a rather mixed response, slightly leaning towards agreeing with the statement.

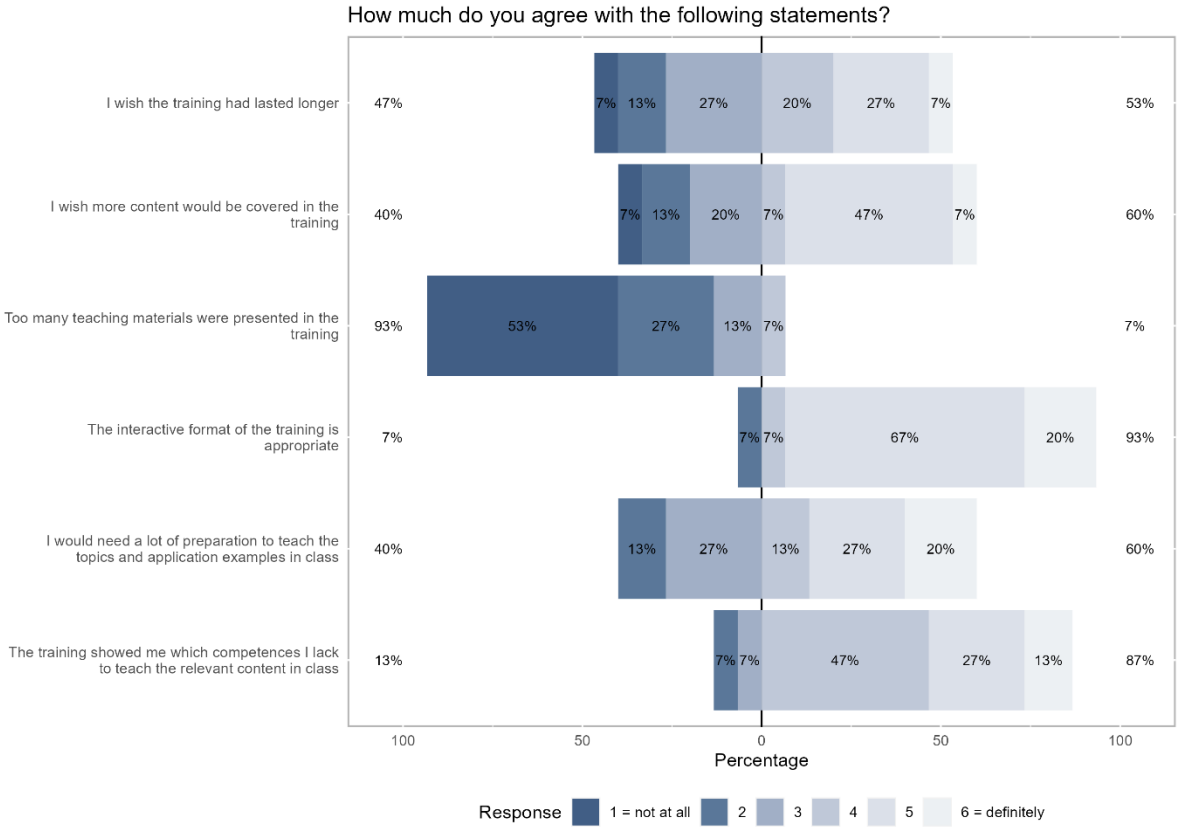


Figure 40 Distribution of the post results for the assessment of the training format and outcomes, 31.01.2023 Vienna, post-evaluation survey, n=15

4.3 Qualitative results

4.3.1 Personal interviews

The themes listed below were deductively derived from the research questions and the structure of the interview guide. The annotations on the right refer to categories that were assigned inductively throughout our qualitative analysis process. Given the novel nature of the research questions, we chose not to limit our analysis to a purely deductive approach. Though our initial research questions and the interview guide shaped the general themes, it was the inductive analysis of participant responses that filled these themes with detailed and meaningful content.

A. Training:

Factors prompting training participation (before the training)

One teacher mentioned attending the TrainDL event, as attending OCG events became a routine.

Participation due to regular attendance of OCG seminars

Other participants mentioned learning about the TrainDL seminar from others: through (former) colleagues or via e-mail/newsletters from the OCG.

Information was obtained through colleagues and professional networks

Expectations (prior to the training)

General expectations are that understanding of the topics will be provided: *"The expectation is, of course, that you still understand everything. That's probably not feasible. But it's my first piece of cake."* (Participant A - 31.01.23, Vienna). The desire to learn.

Building or updating knowledge

Another motivation is expending knowledge, which is evident in the concretely expressed desire for better handling of, for example, data science. Other specific wishes include a foundation for analys-

Need for material

ing tangible problems, solutions that can be immediately implemented in the elective CS course, and a script or textbook containing material on AI

One participant reported having no expectations

No expectations for the training

Design according to previous experience (after the training)

According to the interviewees, the difficulty of the training was adjusted to the levels of prior knowledge in DL and AI: *"In principle, I was led slowly with a bit of explanation."* (Participant F - 31.01.23, Vienna).

Training well-aligned with pre-existing knowledge

Difficulty of the topics for the teachers themselves (after the training)

The topic's complexity varied among respondents, with some finding it especially intricate in areas like AI, while others viewed it as less challenging. However, even those who found it simpler felt a need for deeper engagement: *"If I had to recap the whole thing now and write it together once, I would have to sit down again, read through everything, and work all over again."* (Participant F - 31.01.23, Vienna)

Topics (especially AI) are viewed as rather complex

B. Teaching of DL and AI:

Experience with integration of DL and AI into teaching CS at the secondary level (prior to the training)

Integration can be divided into three areas: (1) lack integration, (2) occasional integration, and (3) integration in the classroom.

(1) The main reason for the lack of integration is the lack of knowledge and available materials that could be easily accessed by

No integration due to the lack of knowledge

the teachers: *"[...] after various definitions are also a bit blurred and everything flows a bit into each other, it is also difficult to really make it clear to the children what exactly artificial intelligence is."* (Participant B - 31.01.23, Vienna)

No integration as integrating DL and AI is not mandatory

Logistical constraints are another factor cited, with some noting that there are simply other priorities to address.

Occasional integration of tools

(2) Topics were also integrated occasionally without an AI focus: *"Unconsciously perhaps, but not focused. So this Google one, I used that of course, but the focus was not on: was that an AI behind it, but on the tool, what can it do? How do I use it?"* (Participant A - 31.01.2023, Vienna)

Integration was implemented via EDLRIS module and with Imaris software

(3) The integration via the EDLRIS (European Driving License for Robots and Intelligent Systems) was positively received. Moreover, analogue exercises rooted in the EDLRIS framework were conducted. Integrations that aren't directly tied to this module were also noted. The Imaris software was utilized effectively as well.

Python was used for ML Examples from everyday life are helpful

In the context of machine learning, Python were successfully used.

It was emphasized that one should compare the exercise used in class to an existing technology in everyday life. *"I would have done it that way, I did it that way. That it is simply clear that we already deal with these techniques naturally. Without us being aware of it. And I think that students need to be a bit more aware of this."* (Participant D - 31.01.23, Vienna)

Ability of (more) effective integration (after training)

This category can be divided into (1) Immediate integration, (2) integration of theory without practical aspects, and (3) integration not yet possible.

(1) Some participants noted that some of the content of the training could be integrated immediately: *"So I already think that I can use the content of today's training, that I can use that. Definitely. Yes, so regarding today, I can take that and copy paste and use it."* (Participant E - 31.01.23, Vienna)

Immediate integration is possible

(2) Other participants feel able to teach theory but not to work practically with examples.

Not ready to teach beyond theory

There is some resistance to using analogue games introduced during the training, hence digital exercises are needed.

Resistance to use analogue exercises, digital exercises are needed

Programs like Orange3 are seen as valuable aids to integration for beginner levels with no programming.

Orange3 can be helpful for beginners

(3) Other interview partners do not see themselves immediately in a position to integrate the training content, since only basic knowledge has been imparted. After all, more than basic knowledge is needed to teach: *"And I think that a teacher probably has to go all the way up so that afterwards he knows how far down he has to go for the (students?)."* (Participant C - 31.01.2023, Vienna)

Training is not sufficient for integration

Given the current content of the training, for successful integration, time-intensive personal efforts are required. These efforts include deeper engagement and rehearsal.

Teachers have to invest time to acquire knowledge and create course content

Difficulty of imparting knowledge to students (after the training)

The following basic requirements for teachability are mentioned: *"I believe that you have to start with the students at a very, very basic level, namely with their own skills. That is, one would have to look at what kind of logical way of thinking students have."* (Participant B - 31.01.23, Vienna)

One should start with students' logical thinking and than move to problem-solving with computers

Afterwards, one can guide students to teach problem-solving with computers or certain programs or to apply them in exercises. Then the step can be completed by a reflection *"if that was your strategy now, how does the computer come up with its strategy?"* (Participant B - 01/31/23, Vienna).

Not all students have basic digital skills required for teaching DL and AI

The relatively new subject "Digital Basic Education" ("Digitale Grundbildung") aims to impart the required digital skills at the lower secondary level, which is a required foundation for teaching DL and AI. However, it's crucial to address the knowledge gap for those students in the upper secondary level who do/did not have access to this subject, ensuring they also reach an appropriate level of understanding.

Orange3 and "Good Monkey - Bad Monkey" game could be exercises with lower threshold

In terms of low-threshold access, Orange3 with its visualization potential could be both accessible and teachable in terms of difficulty. The same applies to analogue exercises such as "Good Monkey - Bad Monkey".

Need more help to set up exercises with Orange3

Some teachers view Orange3 tool to be challenging and need more guidelines on how to set up exercises: *"I already hear students saying: I'm done now. What shall I do next?"* (Participant F - 31.01.23, Vienna)

Orange3 as it is taught within the training requires more concrete guidelines and exercises

There are concerns regarding how students can grasp and internalize the learning content while working independently with Orange3. Nonetheless, it's believed that effective assignments can be crafted using the tool, even if the specifics of these tasks remain uncertain and require additional time.

C. Establishment and steps to integrate DL and AI:

Framework curriculum integration

Regarding its integration into the framework curriculum, the feedback is overwhelmingly positive.

DL and AI are highly relevant for framework curriculum

Incorporating DL and AI into the framework curriculum has been well-received, especially since these technologies are now embedded in numerous applications, we use daily. As DL and AI have a growing impact on society, it's essential that the curriculum emphasizes them, ensuring students can readily relate these subjects to real-world scenarios.

Basic digital education needs to be addressed across all secondary grade levels as a prerequisite for AI integration

Before AI is incorporated into the framework curriculum, there's a suggestion for the stronger integration of digital basic education across all the secondary grades. AI already features in the 'digital basic education' subject at the lower secondary level (secondary level 1). However, it's crucial for it to be consistently addressed across various grade levels. This subject acts as a gateway, paving the way for the future integration of AI and other related topics. It's vital that AI becomes a standard topic from the lower secondary level onwards. As for the broader theme of continuous integration, there's ongoing debate about introducing CS topics at the primary level.

Interdisciplinary integration of AI is needed

Teaching related to the topics should not be isolated to CS but should be understood and integrated as interdisciplinary. This is justified with the consequences of AI: *"Because simply the impact social (unintelligible) thus plays into all areas and that is a social topic simply. And actually all [subjects] would have to make it, which concerns them more or less."* (Participant C - 31.01.2023, Vienna). The topics should be taught for different subjects but with different intensity and depending on

the subject. AI can then be addressed in more detail in the subject of computer science, for example.

CS should be a compulsory subject with at least two hours

For CS, it is considered necessary to make it from an elective to a compulsory subject. In addition, it is currently only planned as a one-year course with two hours in the upper school. This is not sufficient for AI in particular; instead, it should be represented consistently with two hours in the upper school until graduation.

Framework curriculum is already overloaded

Another opinion is that integration into the framework curriculum is challenging, as the framework curriculum is overloaded: *"I don't know. Well, it's so overloaded that when you have to add something new to it, it probably offers little value."* (Participant F - 31.01.23, Vienna)

Integration into the framework curriculum is not essential, as teachers can teach content beyond its scope

In addition, teachers have some autonomy and can decide to teach content beyond the scope of the framework curriculum.

Steps for anchoring in the classroom

Interdisciplinary trainings for CS and teachers from other areas is needed

Consideration is being given to attending advanced training courses, such as the ENARIS (Education and Awareness for Intelligent Systems) workshops, which cater to non-CS teachers as well. Additionally, there's a proposal to provide continuous interdisciplinary training on the topic of AI, collaborating with pedagogical universities.

Integrate DL and AI into the framework curriculum of other subjects

Some teachers think, that DL and AI should be integrated into the framework curriculum of other subjects (non-CS) as well.

Essential measures include building teacher knowledge, allocating appropriate time resources, and incorporating the topics into school textbooks. Beyond just integrating into the general framework curriculum, there's a demand for clear teaching guidelines. As Participant D from Vienna stated on 31.01.23: *"Good learning scenarios that are already worked out. Where even teachers who don't have much of a clue can cope well with it and also have a good introduction."* (Participant D - 31.01.23, Vienna) This would also help to alleviate the apprehensions of teachers less familiar with DL and AI.

Equipping teachers with knowledge, ensuring time resources, incorporating topics into textbooks, and providing clear teaching guidelines

Teacher education

It is logical to include the topics in teacher education or teacher training, otherwise DL and AI cannot be taught at a higher level in school. In university education, AI should be integrated in all or other subjects, but of course at different levels of complexity.

Integrate DL and AI into CS and other subjects within the teacher education

Others see the integration especially useful for the CS subjects and digital basic education; for these subjects, the focus should be broader and include societal aspects as well.

AI should also include societal issues

School curriculum

Computer science classes often lack a set curriculum, with inclusion in the school curriculum largely influenced by individual teacher motivation. The classroom implementation varies, as there isn't always a fixed framework: *"Basically, I think, a lot happens in the classroom life alongside. So it doesn't have to be anchored somewhere. If a teacher is interested in it, likes to do it, they implement it."* (Participant F - 31.01.23, Vienna)

School curriculum is not very important and teachers can easily deviate from it

Barriers

Some believe that there are no inherent barriers to integrating AI into school teaching, noting that teachers have the autonomy to shape their professional routines or collaborate with department colleagues

Autonomy of teachers can make integration possible

Conversely, others anticipate challenges. Among these, a "digital double overload" emerges, where teachers face the dual pressures of adapting to the device initiative [Im Zuge dieser Initiative werden Schülerinnen und Schüler der 5. Schulstufe an teilnehmenden Schulen mit einem Notebook oder Tablet ausgestattet] and introducing AI. Presently, the emphasis is on digitization, with uncertainty about AI's immediate prominence. Time constraints are also highlighted as a standalone barrier.

At the moment AI is not a priority and it can be challenging to integrate it

One challenge is ensuring that topics are presented with the necessary depth, especially in the initial stages like after their inclusion in the framework curriculum. *"Well, I think it's already difficult to bring it into computer science because it's relatively complex and I imagine it's almost more difficult to bring it into the other subjects because it's something completely new, unexpected. So it will certainly take some time for that to take hold."* (Participant C - 01/31/23, Vienna)

Due to complexity, AI cannot be integrated quickly

Some teachers have addressed fears regarding plagiarism through ChatGPT in various subjects and see this as an area that needs guidelines and attention from the school

Fears of ChatGPT-based plagiarism are seen by some as a barrier for AI integration

D. *(Possible) changes through the integration of DL and AI in the framework curricula:*

Students

Student performance test will have to be questioned in the future, but banning tools like the ChatGPT would be the wrong way to go. Students will probably be more careful about what they are allowed to use or not. Teachers, for their part, must of course be sensitized to the fact that students can use AI for their assignments. We must reconsider how to inspire students to find value in learning when AI can easily enhance their grades with a simple command. This shift will likely necessitate a change in the grading system as well.

Academic integrity of students can be a problem when AI tools are pervasive, hence school has to deal with AI

Conversely, some believe that the introduction of such technologies can lead to a deeper understanding of topics, much like the advent of the internet: *" [...] I see it a bit like the introduction of the Internet. There was also something like that, and now I don't need teachers anymore. It's all on the Internet. I don't need to learn anything anymore. And actually, if you formulate it positively, it has led to a deepening [...]."* (Participant A - 31.01.23, Vienna)

AI will not make teachers useless, but can help acquire new knowledge

Regarding AI, its introduction is compared to that of the internet; both present novel subjects demanding substantial knowledge for effective utilization, particularly during initial phases. Furthermore, integrating AI into education ensures students become aware of its existence, functions, and applications. As the future shifts towards automation and AI, fostering an understanding of AI translates to a more realistic and relevant education for students

As automation and AI become more prominent in the future, it's essential for students to understand and engage with these topics

Society

The integration of AI into education will lead to broader societal knowledge and understanding, ultimately driving transformative change within the wider community.: *"Society as a whole, I think if this is already embedded, then it gets into education. Then more and more people know about it. Then, in the long term, something has to change in society. So it will reach the middle of society."* (Participant F - 31.01.23, Vienna)

The integration of AI into education will lead to broader societal knowledge and understanding

Some teachers recognize the significant role AI will have in future workplaces and are concerned about AI replacing human roles.

Concerns regarding AI replacing human roles

E. Training Feedback and Potential for Improvement:

Length

The length is considered to be appropriate and very good for different levels of previous knowledge.

7 hours is a good length

For the average level of previous experience with DL and no experience with AI, one teacher suggested offering a series of shorter modules (around 60 minutes or so per topic) that could be attended online, providing flexibility, but with disadvantages of no personal contacts:

Different length options

"So these long blocks for day-long training sessions are becoming increasingly difficult to integrate into everyday teaching because our calendars are becoming so dense with commitments. And computer scientists often have additional commitments at school. And so getting away is not getting easier, I think it's getting more difficult." (Participant A - 31.01.2023, Vienna)

One-day training is challenging due to teachers workload

Content

The content was perceived as engaging. The games were appreciated in terms of several aspects: *"I also thought the games we did to consolidate basic AI knowledge were great, very originally done, and really educational."* (Participant C - 31.01.2023, Vienna)

Games received positive feedback; engaging content

One suggestion was that the theory part could be expanded, and the practical part shortened.

More theory and less practice

Another suggestion is the inclusion of specific AI topics: *"So I would think what I'm missing a little bit is just AI, Perceptron. The learning, how does forward learning work, backward backpropagation that's what it's called. Exactly. At this level just in Python just practical examples for the computer science teacher."* (Participant E - 31.01.23, Vienna)

Content can include other AI topics with programming (e.g. in Python)

Furthermore, it was wished to move advanced trainings like those of TrainDL more in the direction of programming (like the EDLRIS modules).

Content should contain more programming

Some teachers stated that they need a bridge between analogue exercises and digital tools, like Orange3.

Bridge between analogue and digital exercises was not clear

Opinions on the Orange3 tool are divided; while some express a desire to utilize it more extensively, others find the tool overly abstract and challenging.

Opinions on the Orange3 tool are divided

Opinions on the Orange3 tool are divided

Opinions on the Orange3 tool are divided

TrainDL training (other factors, overall view)

The training was positively received for its expert-led approach and available materials

The training received positive feedback, being described as useful and expert-led. While the content was easy to follow, participants noted a difference between passive listening and active implementation: *"So the degree of difficulty in listening was easy. But that is always one thing. Listening and thinking along and doing it yourself are two big pairs of shoes."* (Participant F - 31.01.23, Vienna) Additionally, participants appreciated the public availability of the materials.

Other trainings (not TrainDL)

There's a strong demand for continuous, up-to-date training on DL and AI, emphasizing the dynamic nature of computer science

There's a call for more frequent and dedicated training on DL and AI due to the evolving nature of computer science. One participant emphasized the importance of continually updating the training content to stay relevant in this dynamic field: *"Accordingly, I would wish to see this more often and for it to be given genuine attention, especially because computer science is such a flexible and diverse sector. It is important to always prioritize the relevance of trainings and ensure an extensive offer for them particularly in this field."* (Participant B - 31.01.23, Vienna) Single workshops or courses aren't enough to capture the topic's complexity. Instead, there should be regular, professionalized training opportunities. Collaboration with the business sector and external institutions is suggested to ensure the training content aligns with professional applications and needs.

4.4 Summary

According to the pre-survey, the participants in the training were in-service CS teachers. The participants' second subjects were diverse, with digital basic education and mathematics being the most popular subjects. Most of the participants had at

least some experience of teaching DL in their classes. The training had a low representation of women.

The findings from the pre- and post-evaluation survey indicate that the training had a positive impact on teachers' perceived competences on how to use DL and AI content in class. The willingness to invest time and effort to incorporate DL and AI was generally quite high. Participants expressed a moderate level of agreement with the statement that they have acquired sufficient competences to teach the learned content in class.

Self-assessment knowledge questions showed that participants improved their understanding of AI concepts. The objective knowledge test showed an improvement for both DL and AI questions, albeit the difference was much more pronounced for the AI part.

All the materials and exercises used in the training were viewed as well suited, with the exercise on data lifecycle using the Orange3 tool receiving lower scores relatively to other examples. With respect to the length of the training, the amount of the covered content as well as the amount of the required preparation, the results are mixed. There is, however, a strong consensus that the training did not have too many teaching materials and that the interactive format was appropriate. Participants, on average, indicate agreement with the statement that the training successfully highlighted the competences they lack to teach the relevant content in classes.

Regarding both DL and AI subjects, the participants tend to agree that the DL and AI content is missing in the existing CS framework curriculum, although there is no solid consensus on this. This viewpoint persists and doesn't notably change post-training. The pre- and post-results also show no significant difference for the statements: "In the future, teaching DL/AI will offer students additional benefits". A post-survey item regarding the societal significance of both topics demonstrated overwhelming agreement that the societal importance of DL and AI is high to warrant their

inclusion in the curriculum. The results also indicate that, on average, teachers anticipate student interest in DL and AI topics and feel confident in their ability to generate excitement among students for these subjects.

Overall, the quantitative findings suggest that the 7-hour CS training in Vienna somewhat enhanced teachers' perceived competences on how to use DL and AI content in class, as well as their understanding of both of these concepts. The results also imply that while participants express positive attitudes towards DL and AI, it cannot be concluded that the training directly influenced these attitudes, as they were measured only post-training. The high level of agreement among participants regarding the training's ability to reveal the competences teachers lack hints at the limitations of a single 7-hour training for the comprehensive integration of DL and AI topics into teaching. Although the self-assessment and knowledge test showed some improvement post-training, on average teachers did not manage to understand all the concepts introduced in class.

The qualitative findings suggest that while the training was structured to cater to varying levels of familiarity with DL and AI, some participants still found the content (especially on AI) to be complex. Some teachers integrated DL and AI at varying levels before attending the training. Those who hadn't taught DL and AI before the training cited a lack of proficiency as the primary reason for not integrating these topics into teaching. Opinions differed on the ease of integrating these topics into teaching, after the training. While some teachers felt to be well-equipped to immediately incorporate some content, others expressed uncertainty and lack of sufficient level of proficiency. It was not clear for some, why both unplugged and digital exercises were introduced and how they fit didactically for which level of school students. The Orange3 tool was controversial, with some praising it (albeit for the levels where no programming can be used), and some finding it challenging,

All teachers unanimously acknowledged the critical importance of AI and the need to impart AI competencies to students. Despite this, the hefty workload of teachers,

combined with their lack of expertise in this domain, presents challenges in introducing these topics, even after a 7-hour training session. It's imperative to consider strategies for providing teachers with comprehensive competencies and to ensure they have the time to participate in such trainings, especially since dedicating even 7 hours can be demanding. Providing specific practical exercises and materials (analogue, digital, with and without programming) tailored to distinct grade levels is crucial. Creating or modifying these resources requires significant time, which teachers often don't have at their disposal.

Teachers mention a number of arguments in favour of the inclusion of DL and AI into the framework curriculum and consequently into teacher education. It is also recommended that DL and AI should taught across various subjects (with a consequence of teaching to non-CS-teachers), with a proposition for CS to be made a mandatory subject. Pertaining to Austria, some teachers believe that before introducing AI into the framework curriculum, there should be a reinforced emphasis on embedding digital basic education across all secondary grades with the sufficient number of hours. One potential challenge for integrating into the framework curriculum is the existing overloaded curriculum. Yet, to implement DL and AI into the classroom, integration into the framework curriculum is mentioned as one of the prerequisites. Another important requirement is appropriate teacher training. In contrast, the school curriculum is seen to be less important.

Teachers voiced concerns about readily accessible Language Learning Models (LLMs), particularly regarding academic integrity. This underscores the need for support with AI not just in terms of teaching methodologies, but also in everyday teaching practices within the classroom and school.

5. CS pre-service training, Germany

5.1 Sample

Out of 25 participants, 18 filled out the post-training evaluation survey. All respondents, except for 2 were between 21 and 36 years old. The share of female respondents was 39%. The respondents reported a wide range of the semesters, from the 1st to the 8th. Most of the respondents (12) leaned towards agreeing with the statement that during their studies they have learnt a lot about DL and AI, while six respondents rather disagreed with this statement.

5.2 Quantitative results

5.2.1 Attitudes towards DL/AI

Table 18 and Figure 41 display the distribution of survey items concerning pre-service teachers' attitudes towards DL/AI. These items were included only in the post-evaluation surveys. The findings reveal overwhelmingly positive attitudes towards incorporating DL and AI into future teaching, highlighting their high societal relevance and pertinence within the context of CS education. Respondents were also asked to rate their interest in integrating DL and AI into their future teaching on a scale ranging from 1 - "not at all" to 6 - "definitely." Out of the 18 respondents, 15 selected either the "5" or "6" category, indicating a very strong interest in integrating DL and AI into their teaching.

Question	Median	Q1	Q3	Min.	Max.	Scale
After the lecture and the seminar on November 2nd and 3rd, my attitude towards incorporating DL and AI in my future teaching is as follows	6.0	6.0	7.0	5	7	1-7
How do you assess the societal relevance of DL and AI?	6.0	5.0	6.0	4	7	1-7
How do you assess the relevance of DL and AI as content in CS education?	5.0	4.25	6.0	3	7	1-7

Table 18 Summary statistics of post-survey results for a set of survey items on pre-service teachers' attitudes towards DL and AI, 02-03.11.2022 Berlin, post-evaluation survey, n=18

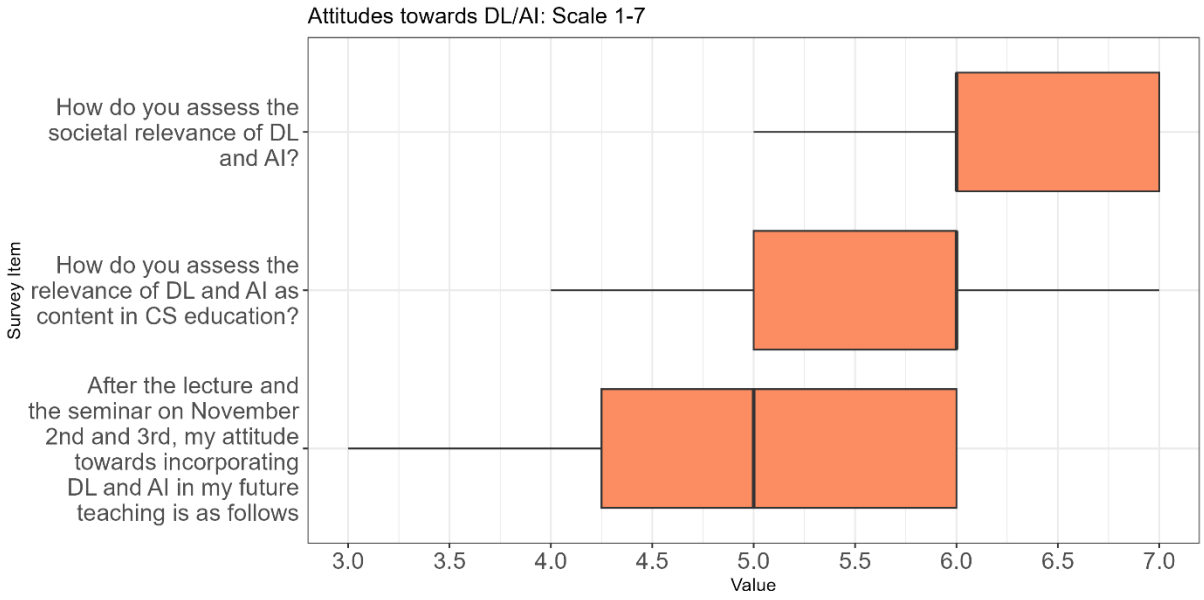


Figure 41 Boxplot comparison of post-survey results for the survey items on pre-service teachers' attitudes towards DL and AI, 02-03.11.2022 Berlin, post-evaluation survey, n=18

5.2.2 Feedback on the learned content and format

Figure 42 showcases pre-service teachers' feedback on the suitability of the training's examples for their future teaching. The results reveal that most participants find the presented examples to be quite suitable for their teaching. Notably, the Orange3

exercise on data lifecycle stood out as one of the most suitable examples. Approximately 89% of respondents rated the Orange3 example with a score of "5" or "6," indicating a high level of suitability.

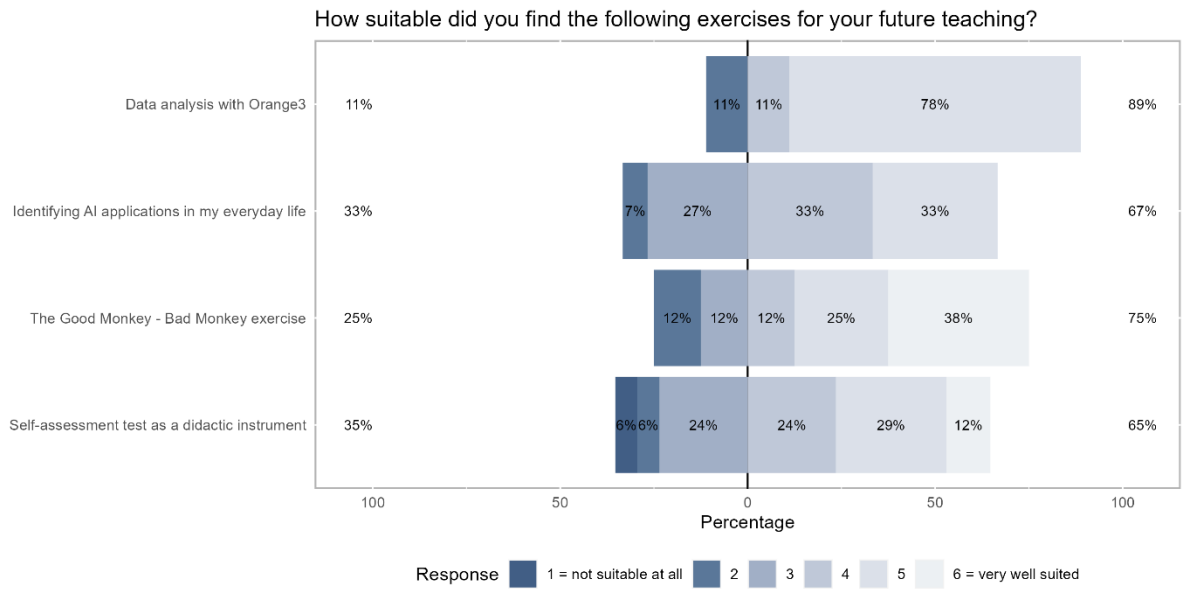


Figure 42 Distribution of the post results for the assessment of the training material/exercises, 02-03.11.2022 Berlin, post-evaluation survey, n=18

5.3 Summary

The quantitative results show overwhelmingly positive attitudes towards incorporating DL and AI into future teaching among the pre-service teachers. The attitudes were assessed using survey items that highlighted the high societal relevance and relevance within the context of CS education. When asked about interest in integrating DL and AI into their future teaching, 15 out of 18 respondents selected either the "5" or "6" category, indicating a strong interest in integration.

Additionally, feedback on the suitability of the training's examples for future teaching was collected. The majority of participants found the presented examples to be quite suitable. Interestingly, while the in-service teachers expressed critical assessments of the Orange 3 tool, the pre-service teachers regarded the Orange3 exercise

on the data lifecycle as a highly suitable example. Approximately 89% of respondents rated it as a "5" or "6" on the suitability scale.

6. Discussion of key findings across the trainings

The pre-evaluation surveys revealed that the three in-service teacher trainings in Germany, Lithuania, and Austria attracted similar participant demographics, albeit with some variations reflective of the countries' wider socio-economic contexts. Almost all attendees were secondary level CS teachers, though a few in addition taught at the primary level. With an exception to a few people, participants were in-service CS teachers teaching a variety of additional subjects, with mathematics being the most common one. While most of the participants had prior experience teaching DL in their classrooms, their exposure to teaching artificial intelligence was comparatively shorter.

Except for one attendee who was completing the practical phase of the training, all interview participants were in-service CS teachers who had voluntarily opted to participate in the interviews.

The majority of respondents at the pre-service workshop were between 21 and 36 years old, with a relatively high representation of female students (39%) compared to the CS trainings in Germany and Austria. The participants reported a wide range of semesters, from the 1st to the 8th. Regarding the perception of learning about DL and AI during their studies, 12 respondents agreed they learned a lot, while six disagreed.

6.1 Quantitative results

All three in-service CS trainings showed that the training had some positive impact on teachers' perceived competences on how to use AI content in class as well as teachers' understanding of AI concepts introduced in the training (see Figure 43, 44, and 45).

For DL, the results were mixed. In Lithuania, there was no significant difference between the pre- and post-training values for the item on how to use DL content in class, albeit the pre-training scores for this item were quite high. While there was an improvement in self-reported understanding of DL concepts in Vilnius and Berlin, Vienna's training showed no statistically significant difference in the pre- and post-training values. The outcomes of the objective knowledge tests paralleled these observations, showcasing more noticeable enhancements in the understanding of AI concepts, as compared to DL, across all the trainings.

Although, on average, teachers in Lithuania obtained higher scores in the self-assessment test for both DL and AI compared to Germany and Austria, they performed relatively worse than the other two respective countries in the objective knowledge test.

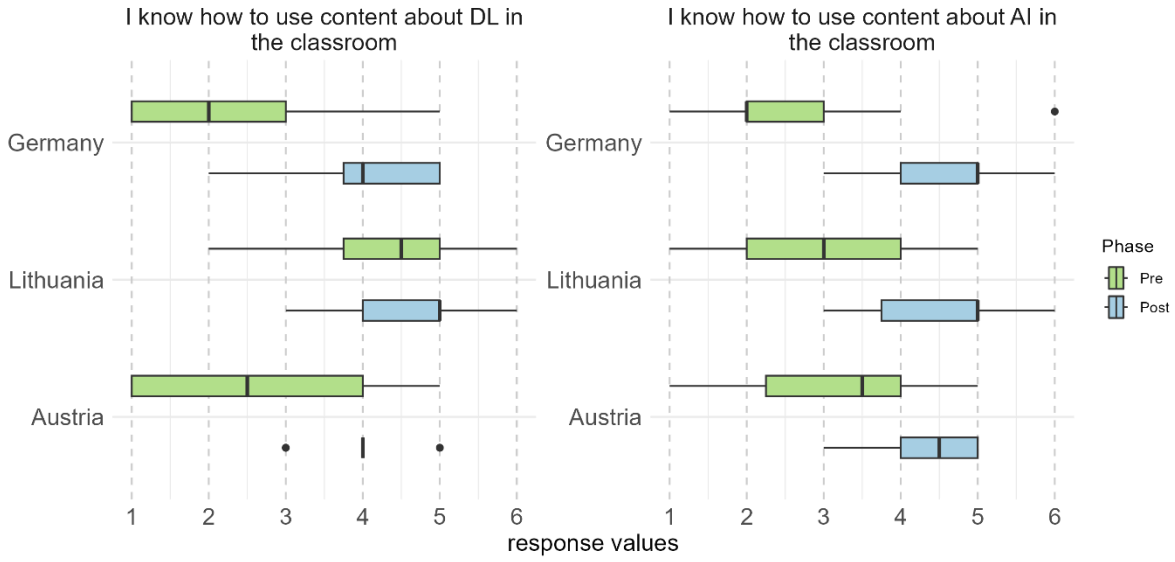


Figure 43 Boxplot comparison of post results for the survey items on teachers' perceived competences to use DL and AI content in class for the three in-service trainings in Germany, Lithuania, and Austria

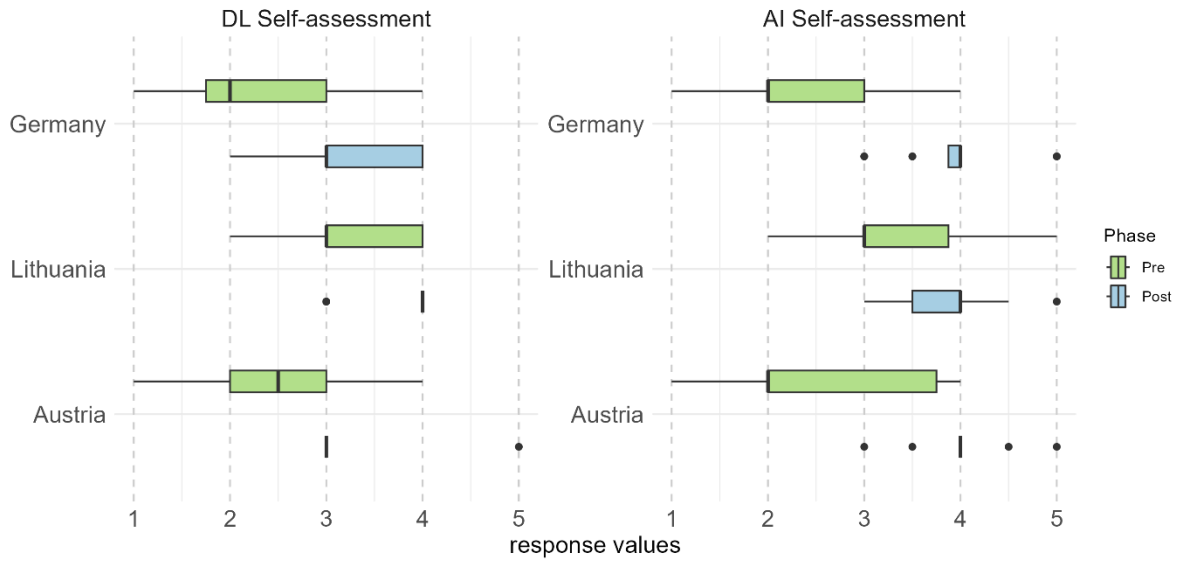


Figure 44 Boxplot comparison of pre- and post-results for self-report knowledge questions on DL (15 questions) and AI (10 questions) for the three in-service trainings in Germany, Lithuania, and Austria

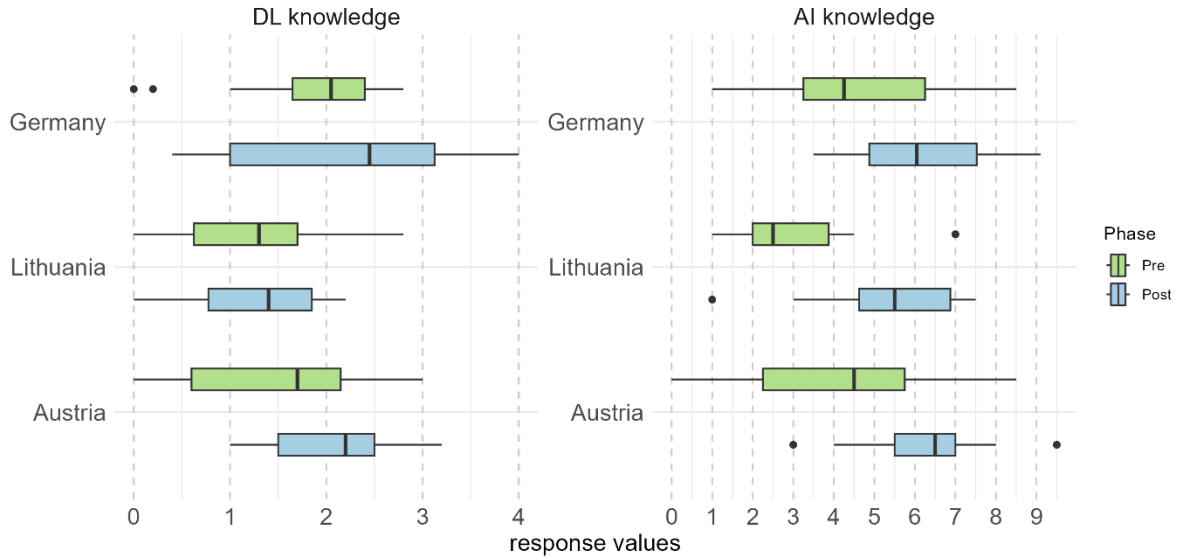


Figure 45 Boxplot comparison of pre- and post-results for knowledge questions on DL (4 questions) and AI (10 questions) for the three in-service trainings in Germany, Lithuania, and Austria

All three locations had the Orange3 tool rated lower regarding its suitability for teaching compared to other exercises and materials, indicating a need for improvements in this area. In Germany and Austria, participants also assessed the topic of data lifecycle as less suitable for their teaching when compared to other topics focusing on AI.

Participants in all three countries tend to agree, that DL and AI content is missing in the current CS curriculum. Furthermore, they agreed on the societal importance of DL and AI, expecting these subjects to generate substantial student interest and engagement.

However, there is not enough evidence to argue, that a single 7-hour training session is sufficient for the comprehensive integration of DL and AI topics into teaching. Overall, while these trainings demonstrated positive impact, they also revealed areas that require improvement and further investigation, such as the delivery of the data lifecycle content. The topic of DL was primarily introduced through the lens of the data lifecycle, which represented a smaller portion of the overall course content. This limited exposure could have restricted the participants' ability to fully grasp the complexity and applicability of DL concepts and the Orange3 tool. The qualitative interviews provided additional insights on the challenges of the DL training content.

For the only pre-service training in Berlin, the quantitative post-survey results show overwhelmingly positive attitudes towards incorporating DL and AI into future teaching among the surveyed pre-service teachers.

6.2 Qualitative results

The qualitative findings across Berlin, Vilnius, and Vienna demonstrate several common themes. Across the three countries, there's a recognized urgency to integrate DL and AI into teaching, given their societal, political, and practical implications. However, doing so requires first of all training teachers in DL and AI.

Participants in the three trainings reported that the training served as a valuable introduction to AI and DL. However, the depth and complexity of certain topics, especially in the realm of AI, were challenging for participants. Exercises involving the Orange3 tool consistently received mixed feedback. While some found it suitable for teaching, others saw it to be rather complex, another critique involved its graphical interface and lack of programming as deterrents for students with a higher level of CS competences.

In all three countries, participants highlighted a gap between training and classroom application. While teachers felt more prepared after the training, the translation of this knowledge into actual lessons would require further engagement with the material. A single 7-hour course is not enough to ensure integration of DL and AI into teaching, especially if teachers did not teach these subjects prior to the training. Teachers expressed a strong desire for concrete course plans and materials tailored to specific grade levels to streamline integration into teaching.

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 framework 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. Most teachers agree that a 7-hour training is not sufficient for proficiency needed to teach DL and AI in class. At the same time, some teachers have difficulties to attend even one-day trainings, due to their teaching responsibilities and high workload. Therefore, the challenge lies in designing a training program that is both

comprehensive and flexible. Such a program could be modular and spread over multiple sessions, allowing teachers to grasp the intricacies of DL and AI at their own pace, without compromising their primary teaching duties.

Teachers voiced concerns about readily accessible Large Language Models (LLMs), particularly regarding academic integrity. This underscores the need for comprehensive AI support, both in terms of teaching methodologies and routine classroom procedures.

6.3 Summary

Findings from our small-n quantitative evaluation suggest that a one-day teacher training may contribute towards an increase in subjectively perceived competences in how to use AI in class as well as in subjectively and objectively assessed content knowledge in the AI-concepts introduced in the training. Improvements in the area of DL focusing on data lifecycle demonstrated high variability with no clear increase post-training. Compared to DL (specifically the topic of data lifecycle using the Orange3 tool), on average AI topics and exercises were perceived to be more suitable for the teaching in the classroom.

In the three countries, teachers acknowledged the significance of integrating DL and AI into the framework curriculum due to their societal and practical implications. Although the trainings provided a foundational understanding of AI and DL, the depth of some topics remained challenging for many. The Orange3 tool evoked mixed responses. In each country, there remained a distinct divide between the training and its classroom implementation. Many believe that a single 7-hour session is insufficient for acquiring enough competences to teach DL and AI in CS classes. Emphasizing the integration of these subjects into the framework curriculums, teachers in Germany and Austria proposed making CS mandatory in secondary education, while some in Lithuania suggested its introduction at the primary level. Given the already full framework curriculums, reprioritizing content is essential. Ultimately, there's a

consensus on the need and urgency of DL and AI teacher training, where the challenge lies in designing a training program that is both comprehensive and flexible and accommodates for teachers packed schedules.

The quantitative results of the pre-service training revealed highly positive attitudes towards incorporating DL and AI into future teaching among the pre-service teachers. Survey items assessing their attitudes displayed very favourable responses, indicating a strong belief in the societal relevance of DL and AI, as well as their significance within the field of CS education. In contrast to the in-service CS teacher, the pre-service respondents saw the data lifecycle examples using a digital tool Orange3 as one of the most suiting examples.

7. Limitations

Some inherent limitations of the evaluation should be considered:

Workshop Format Variation: We did not vary the format and length. Therefore, we cannot trace any impact on the training format as such. Future research could benefit from experimenting with different formats and length.

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 a more robust and generalisable outcome.

Instrument Validity and Reliability: The quality of the survey and knowledge test instruments was not evaluated prior to their application, which could influence the reliability of the data collected. There is an ongoing effort to collect more data and evaluate the questions in future project rounds. All instruments were pre-tested with a small sample of respondents.

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 acquiescence response styles (Rammstedt, Danner, and Bosnjak 2017).

Limited DL Knowledge Test Questions: The DL segment of the study was evaluated using only four knowledge questions, which may not adequately capture the participants' understanding of DL. This limitation could potentially skew the perceived effectiveness of the DL part of the training.

Single Training Sessions: The evaluation was based on a single training session conducted in each of the three countries. This approach may limit the evaluation of consistency in training outcomes across various local contexts within the countries.

Differentiation between subjects: In the interviews and in some survey items, it is not always clear, whether respondents mean their CS subject or their second subject. Although all the trainings were conducted within the context of CS teachers with CS being in the centre of the discussion, one must be careful interpreting the findings.

8. Conclusion

This evaluation report provides insights into the impact of the designed AI and DL training for secondary-level pre- and in-service CS teachers across three European countries – Germany, Lithuania, and Austria. The results point towards the positive influence of these trainings on teacher competences on how to use AI in class and participants' understanding of the AI concepts introduced in the training. The findings for the DL content were mixed, with no clear evidence of improvement in pedagogical and content knowledge. The findings highlight areas for improvement and further research, such as training duration and content depth, as well as the delivery of the DL content. Although sufficient knowledge and materials were provided to offer a

basic introduction to the topics, especially for the AI part, additional training and tailored teaching materials for specified grade levels are necessary to support teachers in developing the required competences for effectively teaching these topics at the secondary level.

The overall positive feedback of the TrainDL training (being seen as a starting point of acquiring competences in DL and AI) coalesced with teachers' positive attitudes towards DL and AI suggest a promising possibility for the integration of DL and AI into the classroom, potentially preparing students for the future societal and occupational landscape. However, the identified limitations of the data from the first round, such as the small sample size and limited variation in the format and content, underline the need for more comprehensive research to refine and optimize similar training programs. Furthermore, the challenges brought to the foreground - such as for example, teacher workload, rapid technological advancements, and limited capacity of the curriculum - offer an opportunity to reflect on the design of future training, curricula and respective policies.

In the forthcoming phases of the TrainDL project and subsequent work in this domain, the following steps are suggested for consideration:

- Further Evaluation of the TrainDL trainings: Continue to collect and analyse (follow-up) data. Given the challenges of small participant groups and hence small sample sizes, the evaluation focus should be on mixed methods approaches. In order to address the challenge of increasing the number of interview participants for the qualitative evaluation given the logistical constraints (it is difficult to schedule multiple individual interviews immediately after the training, as most participants do not have more than ca. 30 minutes available), it is possible to replace individual interviews with focus groups.

- Expanding and Deepening Training Content: Given the mixed results in DL and more positive responses to AI content, the next step can involve expanding the depth and variety of topics covered in these areas. This may include more advanced AI concepts and a broader range of data literacy topics.

- Modular and Flexible Training Programs: To address the challenge of teachers' limited time, future trainings could be designed as modular and blended learning trainings and spread over multiple sessions. This approach would allow teachers to engage with the complexity of DL and AI at their own pace and schedule.

- Tailored Teaching Materials: 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.

- Expanding the Target Audience to non-CS and Primary Teachers: Consider expanding the target audience of the training programs beyond CS teachers to include educators from non-CS subjects as well as primary teachers, reflecting the interdisciplinary nature of DL and AI.

- Addressing Technological Advancements (e.g., LLM): Regularly update training content to keep pace with rapid technological advancements in AI and DL, ensuring that teachers and students are exposed to the latest developments.

- Integrating AI Ethics: Given concerns about Large Language Models and other ethical aspects of AI, future training should include AI ethics and how to integrate this in educational settings.

- Framework Curriculum Integration: Work towards integrating DL and AI into the CS curriculum and possibly into other non-CS subjects. This involves collaborating with educational policymakers and curriculum developers to reprioritise content and make space for these new topics.

- Long-Term Training and Support: Develop long-term training and support mechanisms for teachers, including ongoing professional development opportunities that focus on both pedagogical and technical aspects of DL and AI.

- Building a Community of Practice: Foster a community of practice among educators who are interested in DL and AI. This community could share best practices, resources, and support each other in integrating these topics into their teaching.

Acknowledgment: 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.

9. 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 Interview analyse 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.

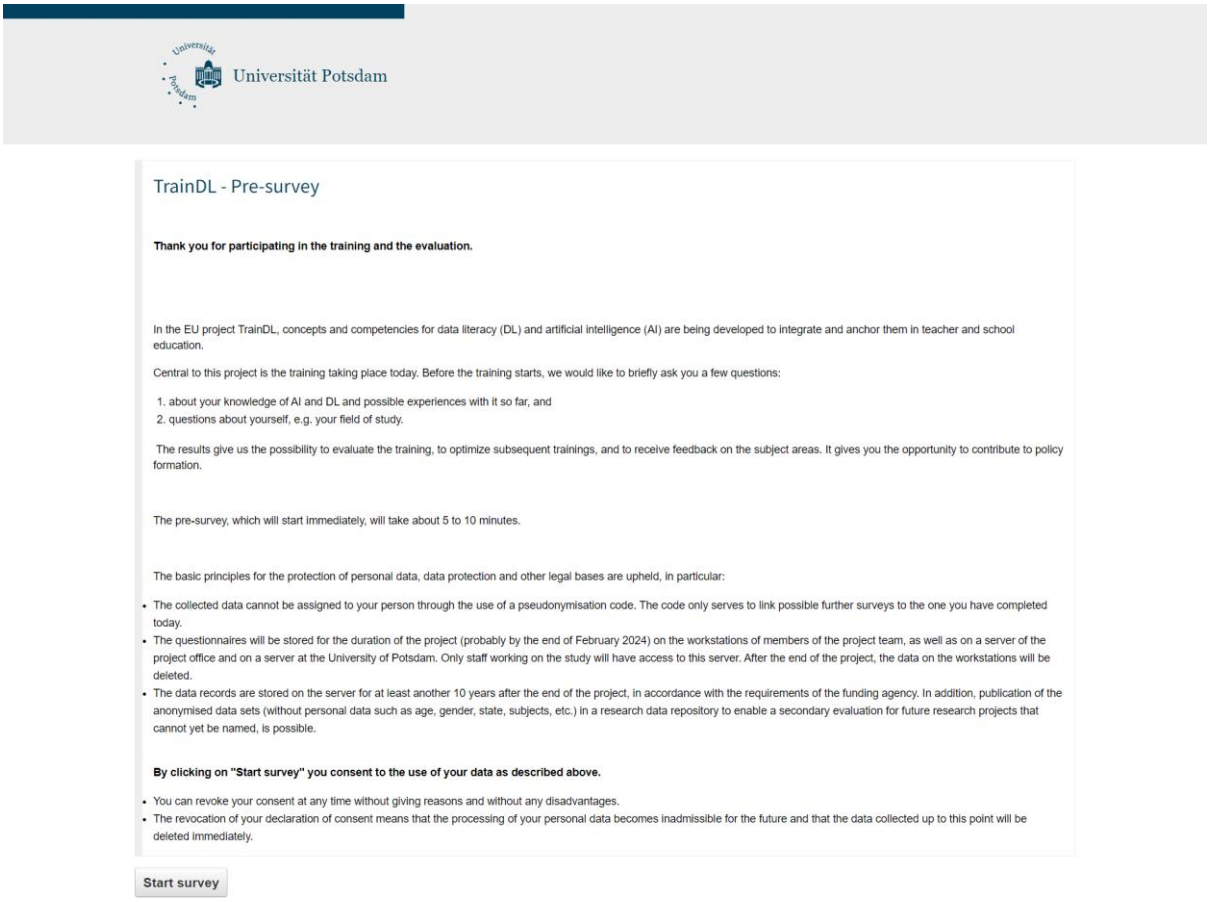
Ridsdale, Chantel, James Rothwell, Mike Smit, Michael Bliemel, Dean Irvine, Dan Kelley, Stan Matwin, Brad Wuetherick, and Hossam ?Ali-Hassan. 2015. "Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report." Accessed July 16, 2023. https://www.researchgate.net/publication/284029915_Strategies_and_Best_Practices_for_Data_Literacy_Educa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

10. Appendix

10.1 CS training, Berlin (13.06.2022): Translated Version of the pre-training evaluation questionnaire



Universität Potsdam

TrainDL - Pre-survey

Thank you for participating in the training and the evaluation.

In the EU project TrainDL, concepts and competencies for data literacy (DL) and artificial intelligence (AI) are being developed to integrate and anchor them in teacher and school education.

Central to this project is the training taking place today. Before the training starts, we would like to briefly ask you a few questions:

1. about your knowledge of AI and DL and possible experiences with it so far, and
2. questions about yourself, e.g. your field of study.

The results give us the possibility to evaluate the training, to optimize subsequent trainings, and to receive feedback on the subject areas. It gives you the opportunity to contribute to policy formation.

The pre-survey, which will start immediately, will take about 5 to 10 minutes.

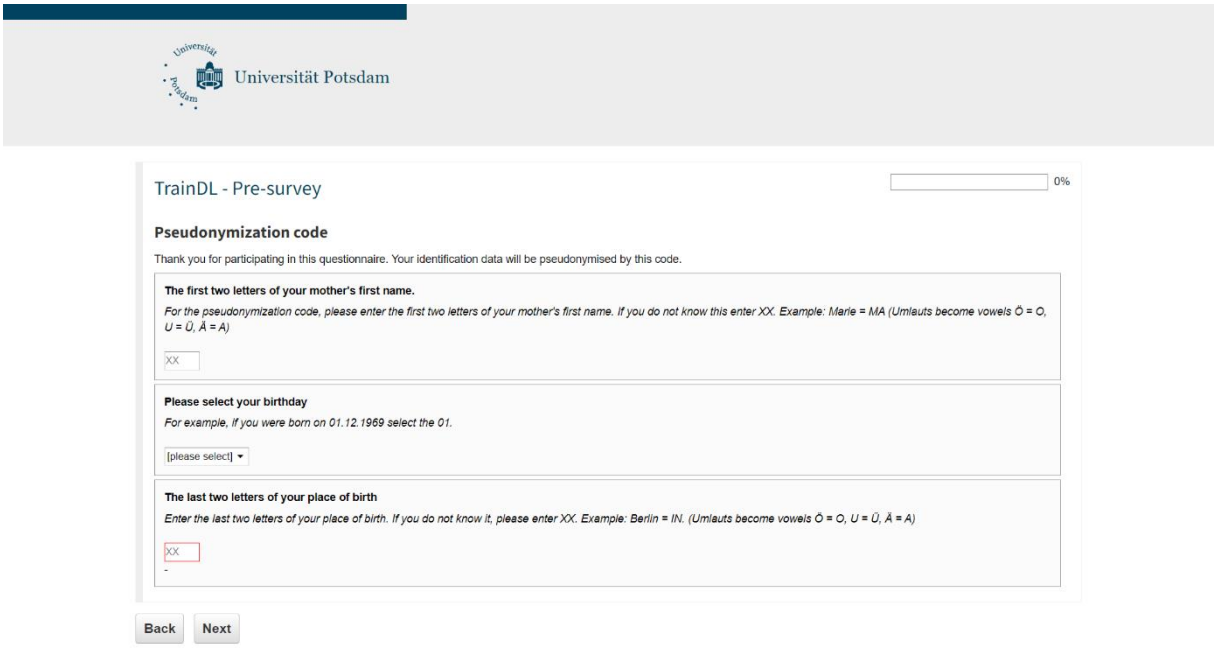
The basic principles for the protection of personal data, data protection and other legal bases are upheld, in particular:

- The collected data cannot be assigned to your person through the use of a pseudonymisation code. The code only serves to link possible further surveys to the one you have completed today.
- The questionnaires will be stored for the duration of the project (probably by the end of February 2024) on the workstations of members of the project team, as well as on a server of the project office and on a server at the University of Potsdam. Only staff working on the study will have access to this server. After the end of the project, the data on the workstations will be deleted.
- The data records are stored on the server for at least another 10 years after the end of the project, in accordance with the requirements of the funding agency. In addition, publication of the anonymised data sets (without personal data such as age, gender, state, subjects, etc.) in a research data repository to enable a secondary evaluation for future research projects that cannot yet be named, is possible.

By clicking on "Start survey" you consent to the use of your data as described above.

- You can revoke your consent at any time without giving reasons and without any disadvantages.
- The revocation of your declaration of consent means that the processing of your personal data becomes inadmissible for the future and that the data collected up to this point will be deleted immediately.

Start survey



Universität Potsdam

TrainDL - Pre-survey

0%

Pseudonymization code

Thank you for participating in this questionnaire. Your identification data will be pseudonymised by this code.

The first two letters of your mother's first name.

For the pseudonymization code, please enter the first two letters of your mother's first name. If you do not know this enter XX. Example: Marie = MA (Umlauts become vowels Ö = O, U = Ü, Ä = A)

XX

Please select your birthday

For example, If you were born on 01.12.1969 select the 01.

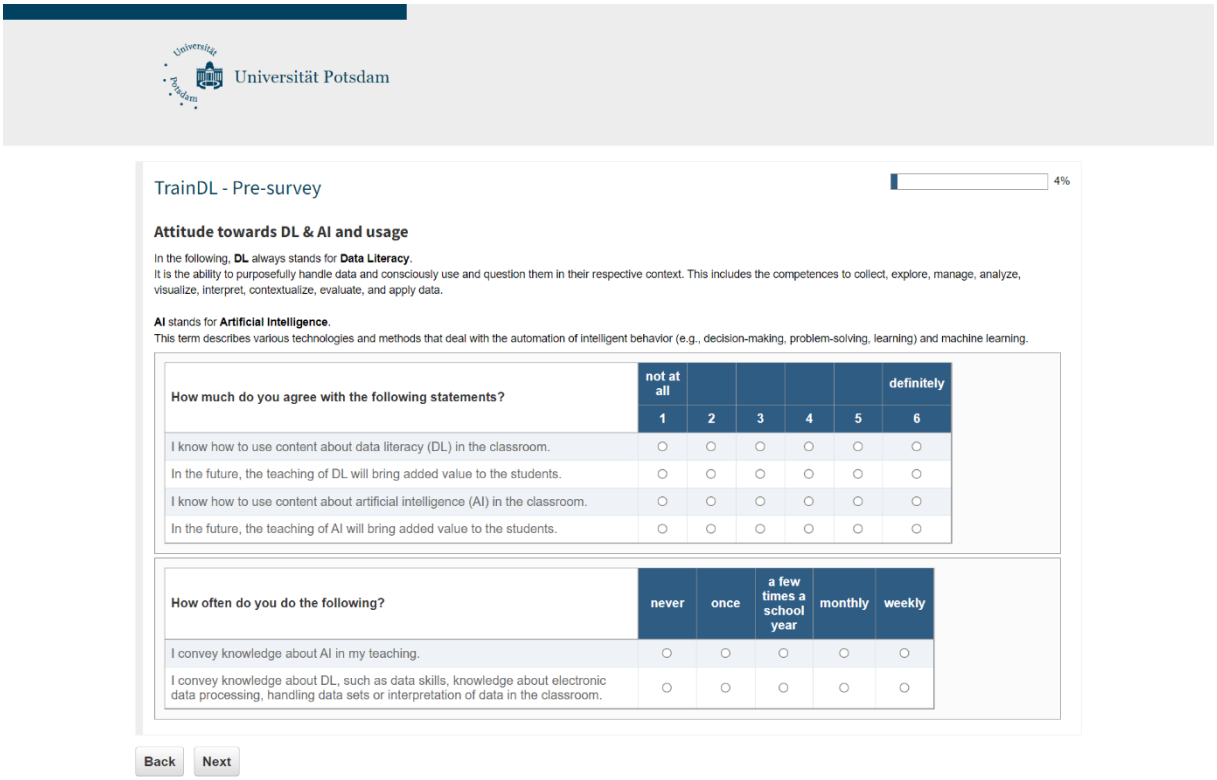
[please select]

The last two letters of your place of birth

Enter the last two letters of your place of birth. If you do not know it, please enter XX. Example: Berlin = IN. (Umlauts become vowels Ö = O, U = Ü, Ä = A)

XX

Back Next



Universität Potsdam

TrainDL - Pre-survey

4%

Attitude towards DL & AI and usage

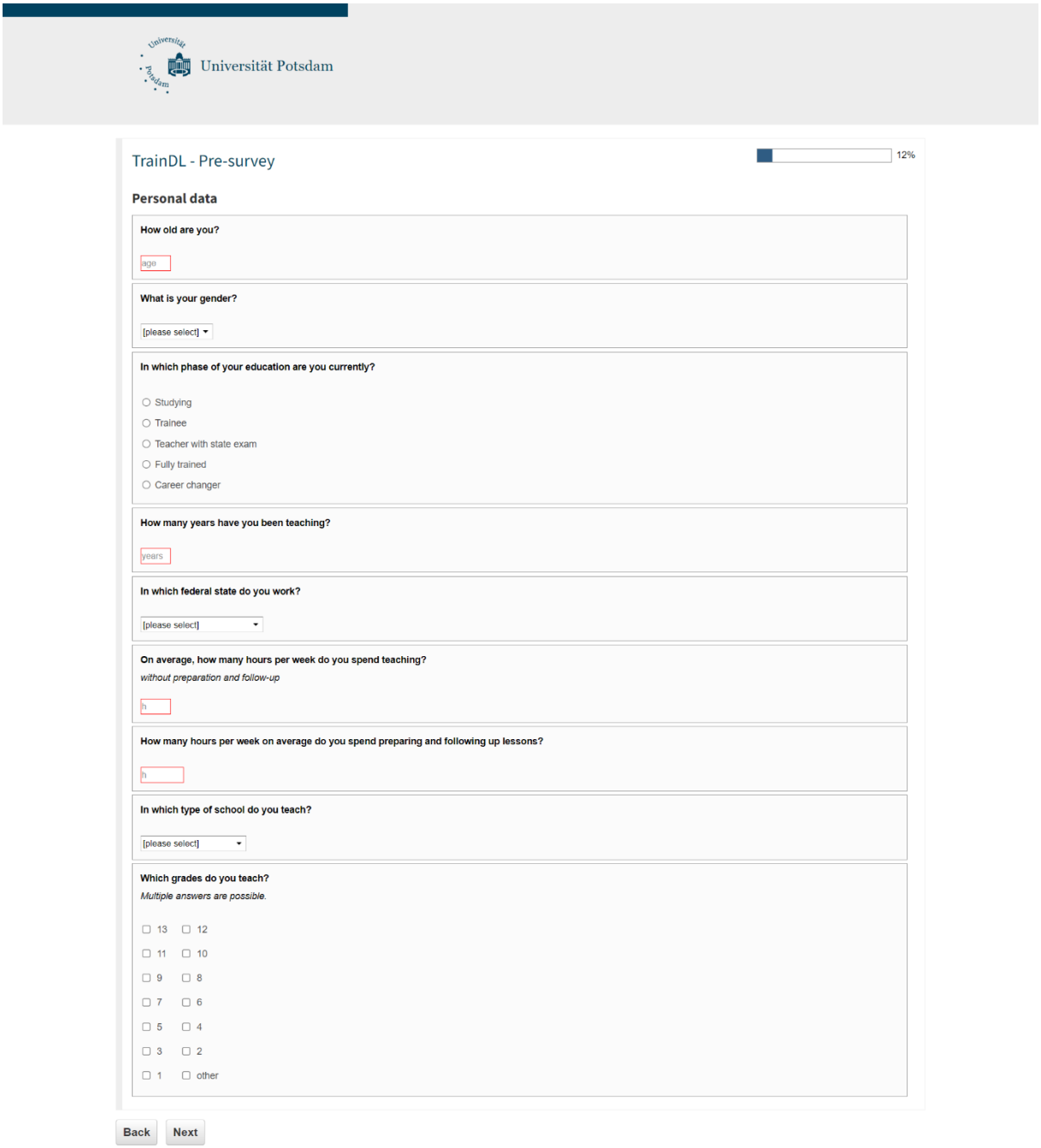
In the following, **DL** always stands for **Data Literacy**.
It is the ability to purposefully handle data and consciously use and question them in their respective context. This includes the competences to collect, explore, manage, analyze, visualize, interpret, contextualize, evaluate, and apply data.

AI stands for **Artificial Intelligence**.
This term describes various technologies and methods that deal with the automation of intelligent behavior (e.g., decision-making, problem-solving, learning) and machine learning.

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I know how to use content about data literacy (DL) in the classroom.	○	○	○	○	○	○
In the future, the teaching of DL will bring added value to the students.	○	○	○	○	○	○
I know how to use content about artificial intelligence (AI) in the classroom.	○	○	○	○	○	○
In the future, the teaching of AI will bring added value to the students.	○	○	○	○	○	○

How often do you do the following?	never	once	a few times a school year	monthly	weekly
I convey knowledge about AI in my teaching.	○	○	○	○	○
I convey knowledge about DL, such as data skills, knowledge about electronic data processing, handling data sets or interpretation of data in the classroom.	○	○	○	○	○

Back Next



Universität Potsdam

TrainDL - Pre-survey

12%

Personal data

How old are you?

age

What is your gender?

[please select]

In which phase of your education are you currently?

- Studying
- Trainee
- Teacher with state exam
- Fully trained
- Career changer

How many years have you been teaching?

years

In which federal state do you work?

[please select]

On average, how many hours per week do you spend teaching?

without preparation and follow-up

h

How many hours per week on average do you spend preparing and following up lessons?

h

In which type of school do you teach?

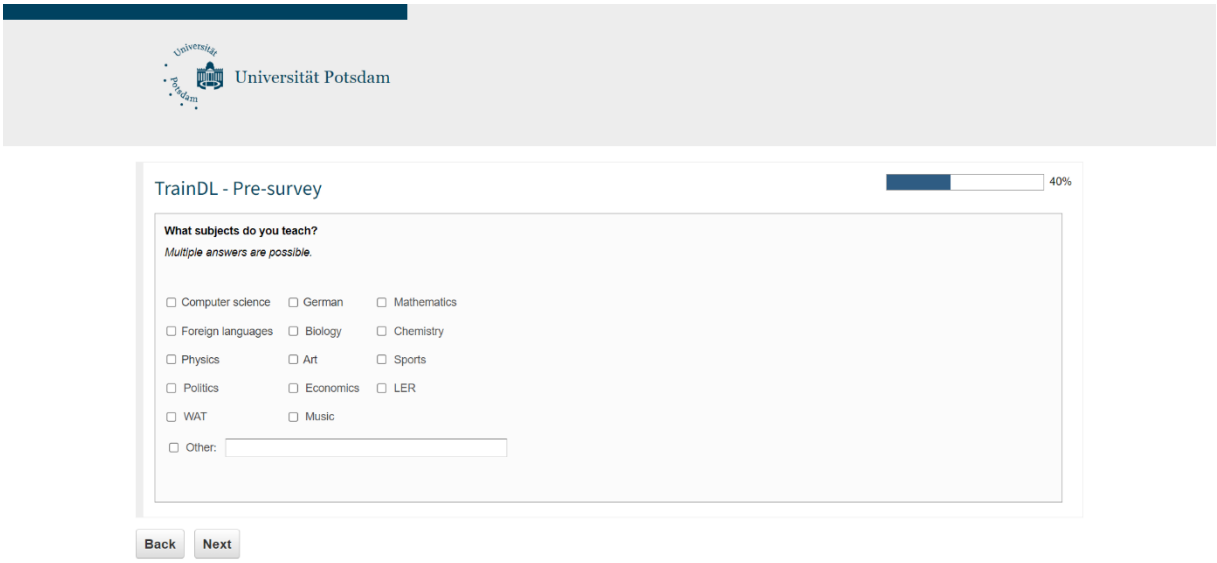
[please select]

Which grades do you teach?

Multiple answers are possible.

- 13
- 12
- 11
- 10
- 9
- 8
- 7
- 6
- 5
- 4
- 3
- 2
- 1
- other

Back Next



Universität Potsdam

TrainDL - Pre-survey

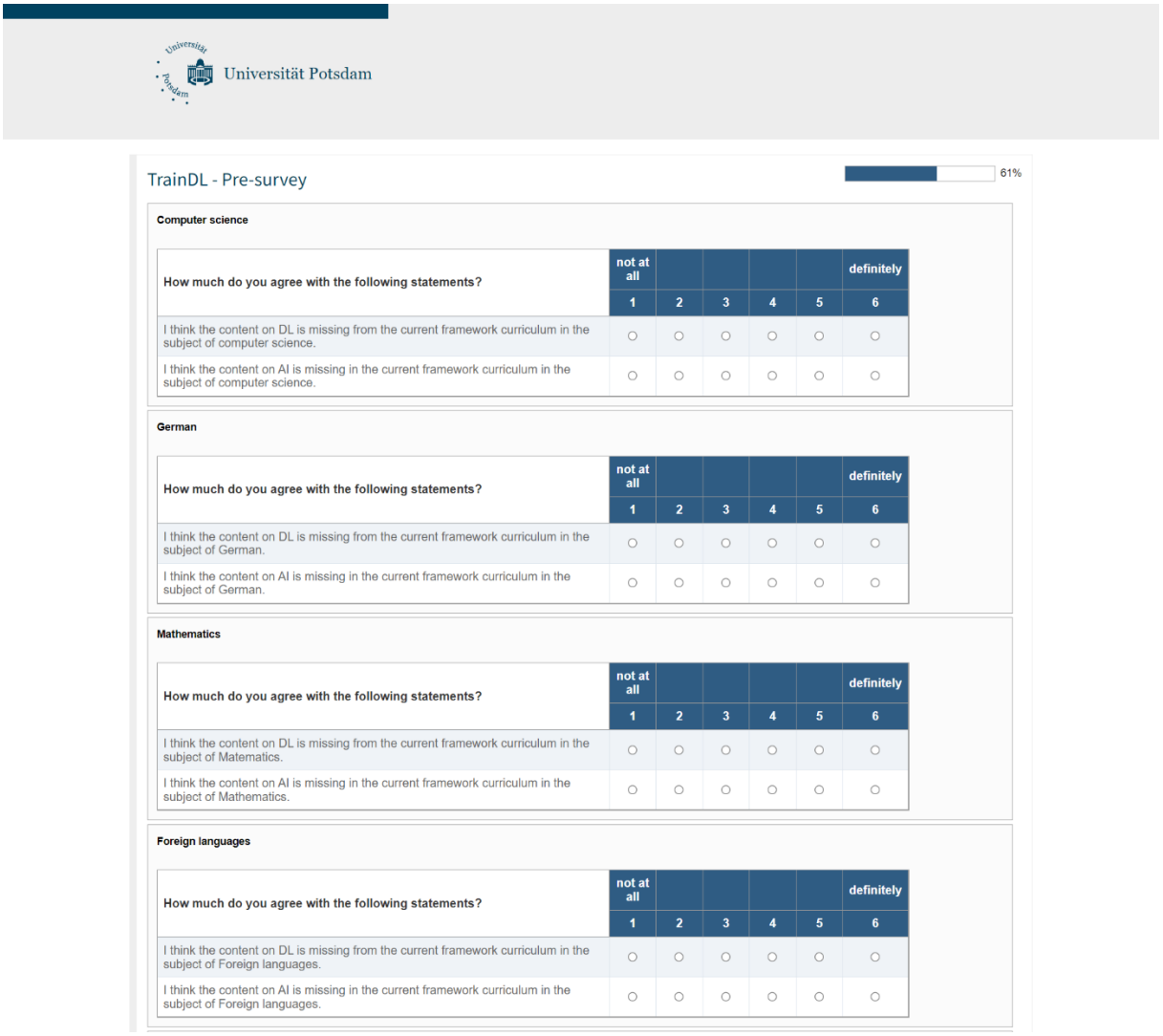
40%

What subjects do you teach?

Multiple answers are possible.

- ☐ Computer science
- ☐ German
- ☐ Mathematics
- ☐ Foreign languages
- ☐ Biology
- ☐ Chemistry
- ☐ Physics
- ☐ Art
- ☐ Sports
- ☐ Politics
- ☐ Economics
- ☐ LER
- ☐ WAT
- ☐ Music
- ☐ Other:

Back Next



Universität Potsdam

TrainDL - Pre-survey

61%

Computer science

How much do you agree with the following statements?	not at all 1	2	3	4	5	definitely 6
I think the content on DL is missing from the current framework curriculum in the subject of computer science.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of computer science.	○	○	○	○	○	○

German

How much do you agree with the following statements?	not at all 1	2	3	4	5	definitely 6
I think the content on DL is missing from the current framework curriculum in the subject of German.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of German.	○	○	○	○	○	○

Mathematics

How much do you agree with the following statements?	not at all 1	2	3	4	5	definitely 6
I think the content on DL is missing from the current framework curriculum in the subject of Matematics.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of Mathematics.	○	○	○	○	○	○

Foreign languages

How much do you agree with the following statements?	not at all 1	2	3	4	5	definitely 6
I think the content on DL is missing from the current framework curriculum in the subject of Foreign languages.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of Foreign languages.	○	○	○	○	○	○

Biology

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of biology.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of biology.	○	○	○	○	○	○

Chemistry

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of chemics.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of chemics.	○	○	○	○	○	○

Physics

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of physics.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of physics.	○	○	○	○	○	○

Art

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of art.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of art.	○	○	○	○	○	○

Sports

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of sports.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of sports.	○	○	○	○	○	○

Politics

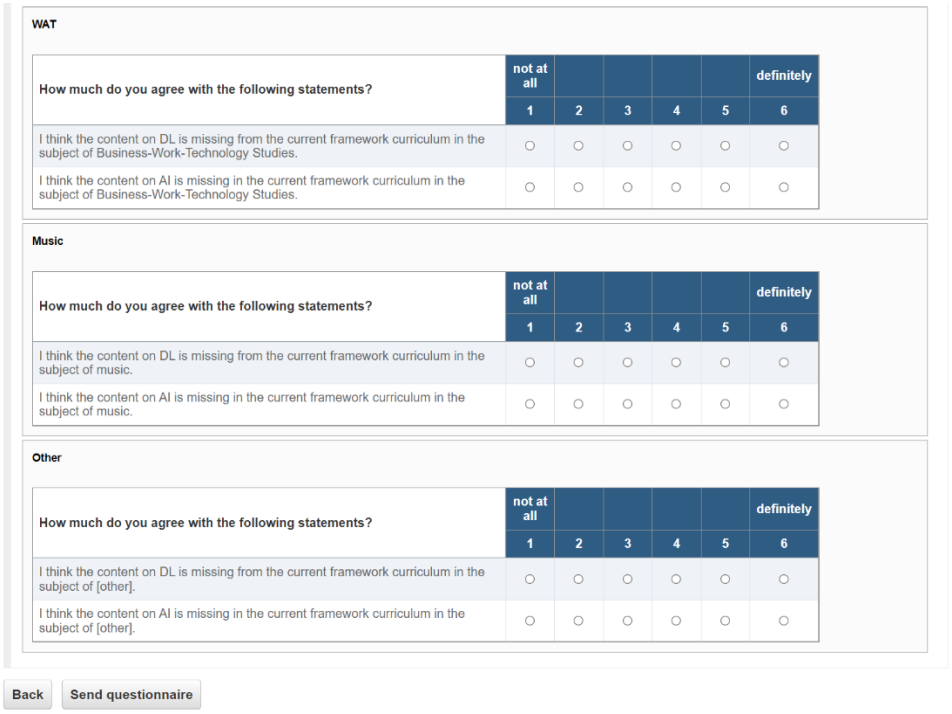
How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of politics.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of politics.	○	○	○	○	○	○

Economics

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of economics.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of economics.	○	○	○	○	○	○

LER

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Life-Design-Ethics-Religious Studies.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of Life-Design-Ethics-Religious Studies.	○	○	○	○	○	○



WAT

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Business-Work-Technology Studies.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of Business-Work-Technology Studies.	○	○	○	○	○	○

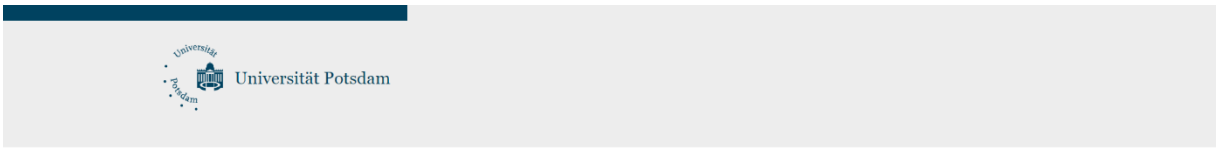
Music

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of music.	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of music.	○	○	○	○	○	○

Other

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of [other].	○	○	○	○	○	○
I think the content on AI is missing in the current framework curriculum in the subject of [other].	○	○	○	○	○	○

Back | Send questionnaire



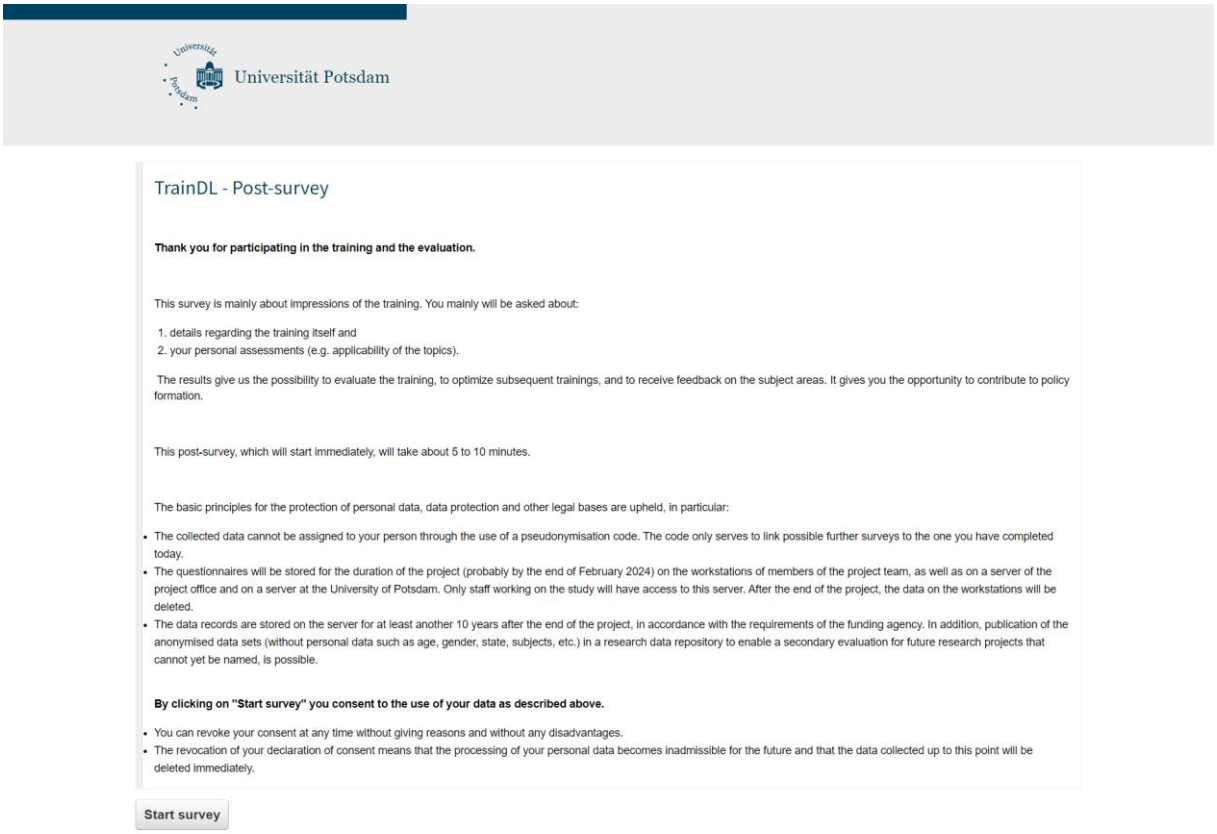
TrainDL - Pre-survey

Thank you for your participation.

To exit properly, please click on "Finish survey"

Finish survey

10.2 CS training, Berlin (13.06.2022): Translated Version of the post-training evaluation questionnaire



Universität Potsdam

TrainDL - Post-survey

Thank you for participating in the training and the evaluation.

This survey is mainly about impressions of the training. You mainly will be asked about:

1. details regarding the training itself and
2. your personal assessments (e.g. applicability of the topics).

The results give us the possibility to evaluate the training, to optimize subsequent trainings, and to receive feedback on the subject areas. It gives you the opportunity to contribute to policy formation.

This post-survey, which will start immediately, will take about 5 to 10 minutes.

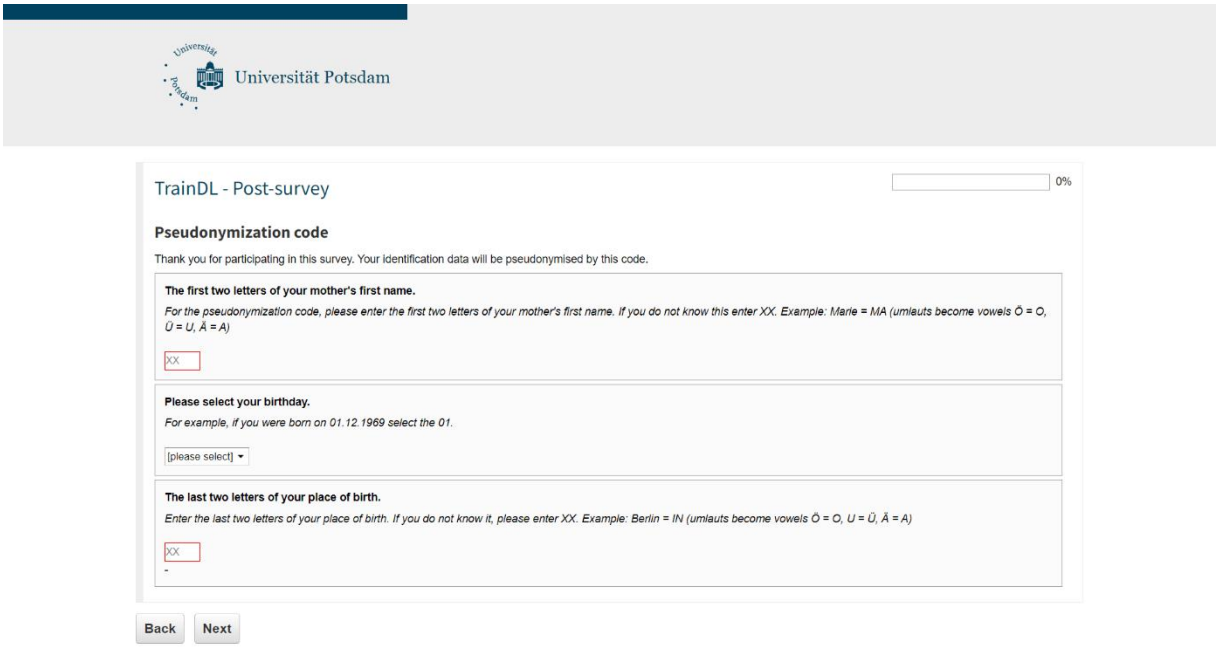
The basic principles for the protection of personal data, data protection and other legal bases are upheld, in particular:

- The collected data cannot be assigned to your person through the use of a pseudonymisation code. The code only serves to link possible further surveys to the one you have completed today.
- The questionnaires will be stored for the duration of the project (probably by the end of February 2024) on the workstations of members of the project team, as well as on a server of the project office and on a server at the University of Potsdam. Only staff working on the study will have access to this server. After the end of the project, the data on the workstations will be deleted.
- The data records are stored on the server for at least another 10 years after the end of the project, in accordance with the requirements of the funding agency. In addition, publication of the anonymised data sets (without personal data such as age, gender, state, subjects, etc.) in a research data repository to enable a secondary evaluation for future research projects that cannot yet be named, is possible.

By clicking on "Start survey" you consent to the use of your data as described above.

- You can revoke your consent at any time without giving reasons and without any disadvantages.
- The revocation of your declaration of consent means that the processing of your personal data becomes inadmissible for the future and that the data collected up to this point will be deleted immediately.

Start survey



Universität Potsdam

TrainDL - Post-survey

0%

Pseudonymization code

Thank you for participating in this survey. Your identification data will be pseudonymised by this code.

The first two letters of your mother's first name.

For the pseudonymization code, please enter the first two letters of your mother's first name. If you do not know this enter XX. Example: Marie = MA (umlauts become vowels Ö = O, Ü = U, Ä = A)

XX

Please select your birthday.

For example, if you were born on 01.12.1969 select the 01.

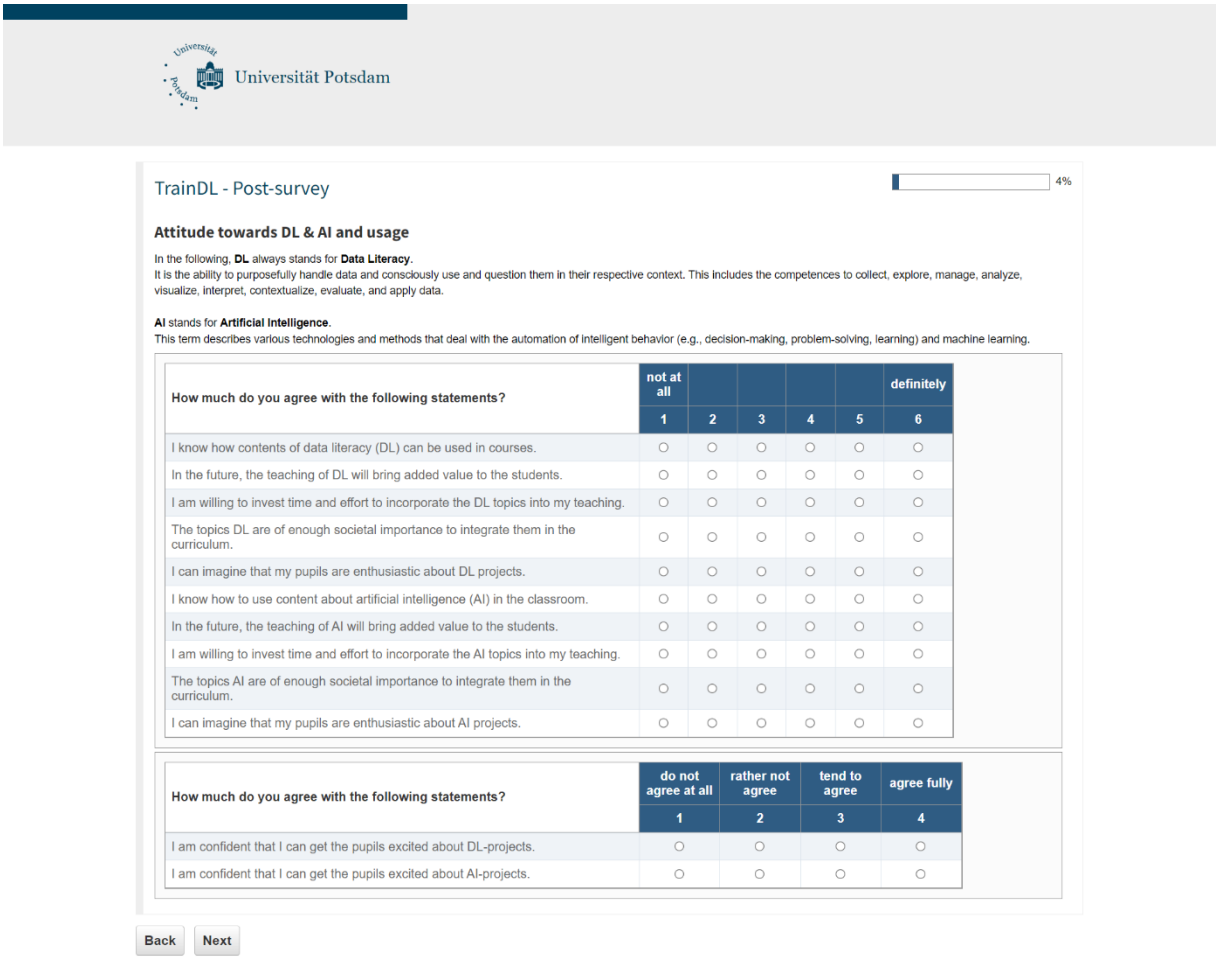
[please select]

The last two letters of your place of birth.

Enter the last two letters of your place of birth. If you do not know it, please enter XX. Example: Berlin = IN (umlauts become vowels Ö = O, U = Ü, Ä = A)

XX

Back Next



Universität Potsdam

TrainDL - Post-survey

4%

Attitude towards DL & AI and usage

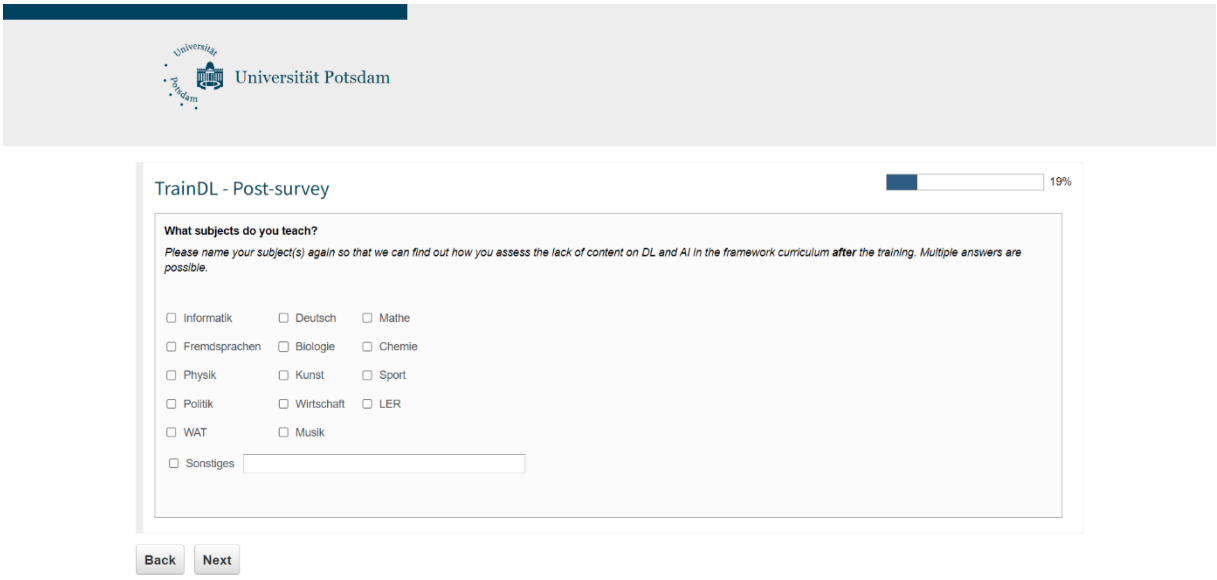
In the following, **DL** always stands for **Data Literacy**.
It is the ability to purposefully handle data and consciously use and question them in their respective context. This includes the competences to collect, explore, manage, analyze, visualize, interpret, contextualize, evaluate, and apply data.

AI stands for **Artificial Intelligence**.
This term describes various technologies and methods that deal with the automation of intelligent behavior (e.g., decision-making, problem-solving, learning) and machine learning.

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I know how contents of data literacy (DL) can be used in courses.	○	○	○	○	○	○
In the future, the teaching of DL will bring added value to the students.	○	○	○	○	○	○
I am willing to invest time and effort to incorporate the DL topics into my teaching.	○	○	○	○	○	○
The topics DL are of enough societal importance to integrate them in the curriculum.	○	○	○	○	○	○
I can imagine that my pupils are enthusiastic about DL projects.	○	○	○	○	○	○
I know how to use content about artificial intelligence (AI) in the classroom.	○	○	○	○	○	○
In the future, the teaching of AI will bring added value to the students.	○	○	○	○	○	○
I am willing to invest time and effort to incorporate the AI topics into my teaching.	○	○	○	○	○	○
The topics AI are of enough societal importance to integrate them in the curriculum.	○	○	○	○	○	○
I can imagine that my pupils are enthusiastic about AI projects.	○	○	○	○	○	○

How much do you agree with the following statements?	do not agree at all	rather not agree	tend to agree	agree fully
	1	2	3	4
I am confident that I can get the pupils excited about DL-projects.	○	○	○	○
I am confident that I can get the pupils excited about AI-projects.	○	○	○	○

Back Next



Universität Potsdam

TrainDL - Post-survey

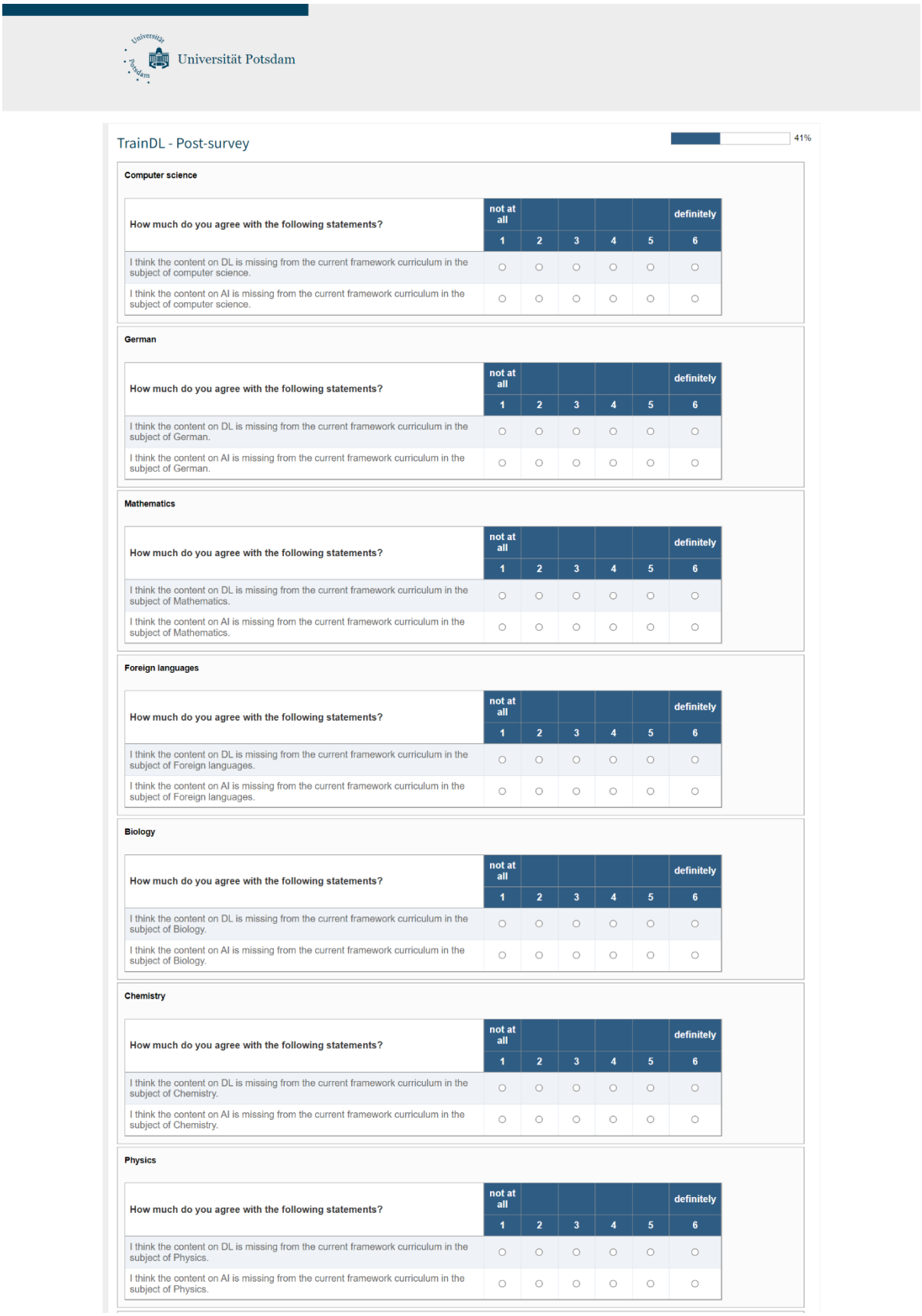
19%

What subjects do you teach?

Please name your subject(s) again so that we can find out how you assess the lack of content on DL and AI in the framework curriculum ***after*** *the training. Multiple answers are possible.*

☐ Informatik ☐ Deutsch ☐ Mathe
☐ Fremdsprachen ☐ Biologie ☐ Chemie
☐ Physik ☐ Kunst ☐ Sport
☐ Politik ☐ Wirtschaft ☐ LER
☐ WAT ☐ Musik
☐ Sonstiges

Back Next



Universität Potsdam

TrainDL - Post-survey

41%

Computer science

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of computer science.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of computer science.	○	○	○	○	○	○

German

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of German.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of German.	○	○	○	○	○	○

Mathematics

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Mathematics.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Mathematics.	○	○	○	○	○	○

Foreign languages

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Foreign languages.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Foreign languages.	○	○	○	○	○	○

Biology

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Biology.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Biology.	○	○	○	○	○	○

Chemistry

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Chemistry.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Chemistry.	○	○	○	○	○	○

Physics

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Physics.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Physics.	○	○	○	○	○	○

Art

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Art.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Art.	○	○	○	○	○	○

Sports

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Sports.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Sports.	○	○	○	○	○	○

Politics

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Politics.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Politics.	○	○	○	○	○	○

Economics

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Economics.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Economics.	○	○	○	○	○	○

Life-Design-Ethics-Religious Studies

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Life-Design-Ethics-Religious Studies.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Life-Design-Ethics-Religious Studies.	○	○	○	○	○	○

Business-Work-Technology Studies

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Business-Work-Technology Studies.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Business-Work-Technology Studies.	○	○	○	○	○	○

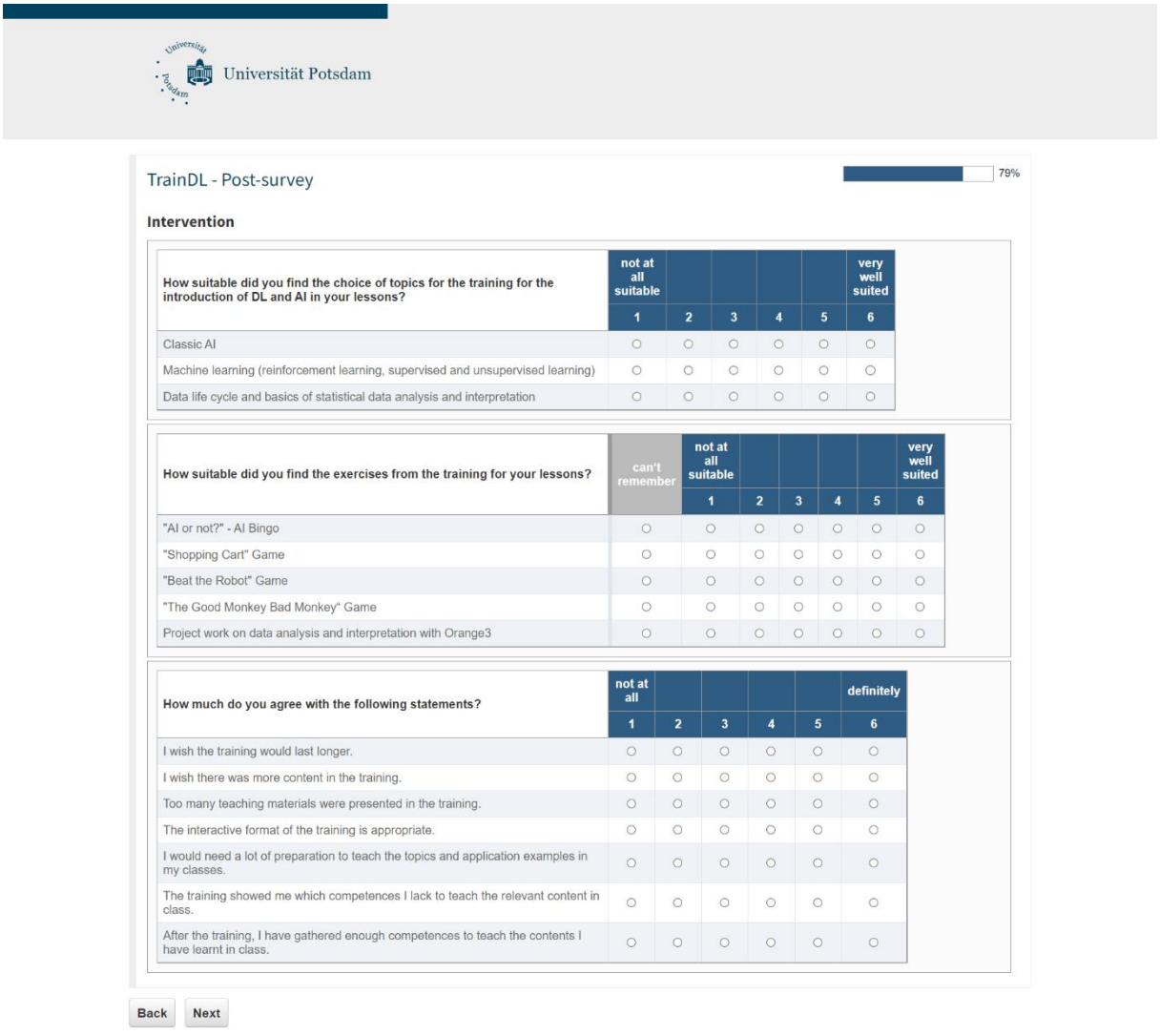
Music

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of Music.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of Music.	○	○	○	○	○	○

Other

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I think the content on DL is missing from the current framework curriculum in the subject of not asked.	○	○	○	○	○	○
I think the content on AI is missing from the current framework curriculum in the subject of not asked.	○	○	○	○	○	○

Back Next



Universität Potsdam

TrainDL - Post-survey

79%

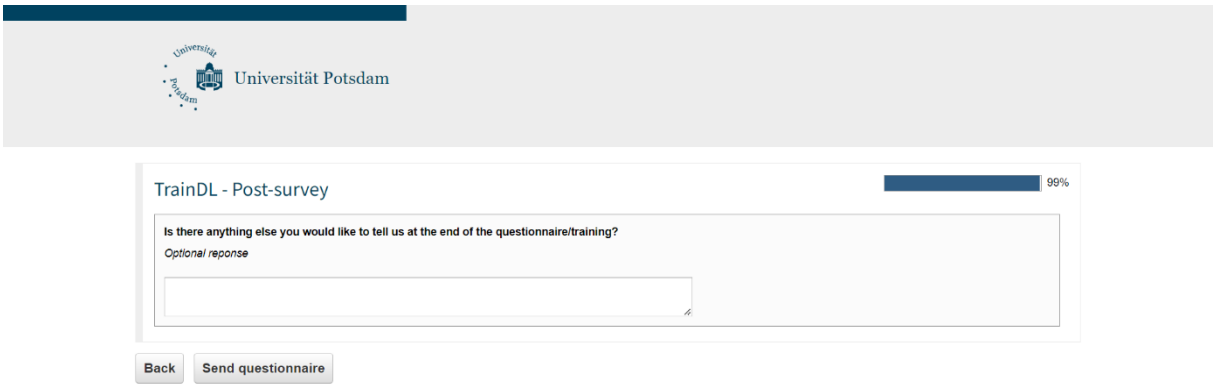
Intervention

How suitable did you find the choice of topics for the training for the introduction of DL and AI in your lessons?	not at all suitable					very well suited
	1	2	3	4	5	6
Classic AI	○	○	○	○	○	○
Machine learning (reinforcement learning, supervised and unsupervised learning)	○	○	○	○	○	○
Data life cycle and basics of statistical data analysis and interpretation	○	○	○	○	○	○

How suitable did you find the exercises from the training for your lessons?	can't remember	not at all suitable					very well suited
		1	2	3	4	5	6
"AI or not?" - AI Bingo	○	○	○	○	○	○	○
"Shopping Cart" Game	○	○	○	○	○	○	○
"Beat the Robot" Game	○	○	○	○	○	○	○
"The Good Monkey Bad Monkey" Game	○	○	○	○	○	○	○
Project work on data analysis and interpretation with Orange3	○	○	○	○	○	○	○

How much do you agree with the following statements?	not at all					definitely
	1	2	3	4	5	6
I wish the training would last longer.	○	○	○	○	○	○
I wish there was more content in the training.	○	○	○	○	○	○
Too many teaching materials were presented in the training.	○	○	○	○	○	○
The interactive format of the training is appropriate.	○	○	○	○	○	○
I would need a lot of preparation to teach the topics and application examples in my classes.	○	○	○	○	○	○
The training showed me which competences I lack to teach the relevant content in class.	○	○	○	○	○	○
After the training, I have gathered enough competences to teach the contents I have learnt in class.	○	○	○	○	○	○

Back Next



Universität Potsdam

TrainDL - Post-survey

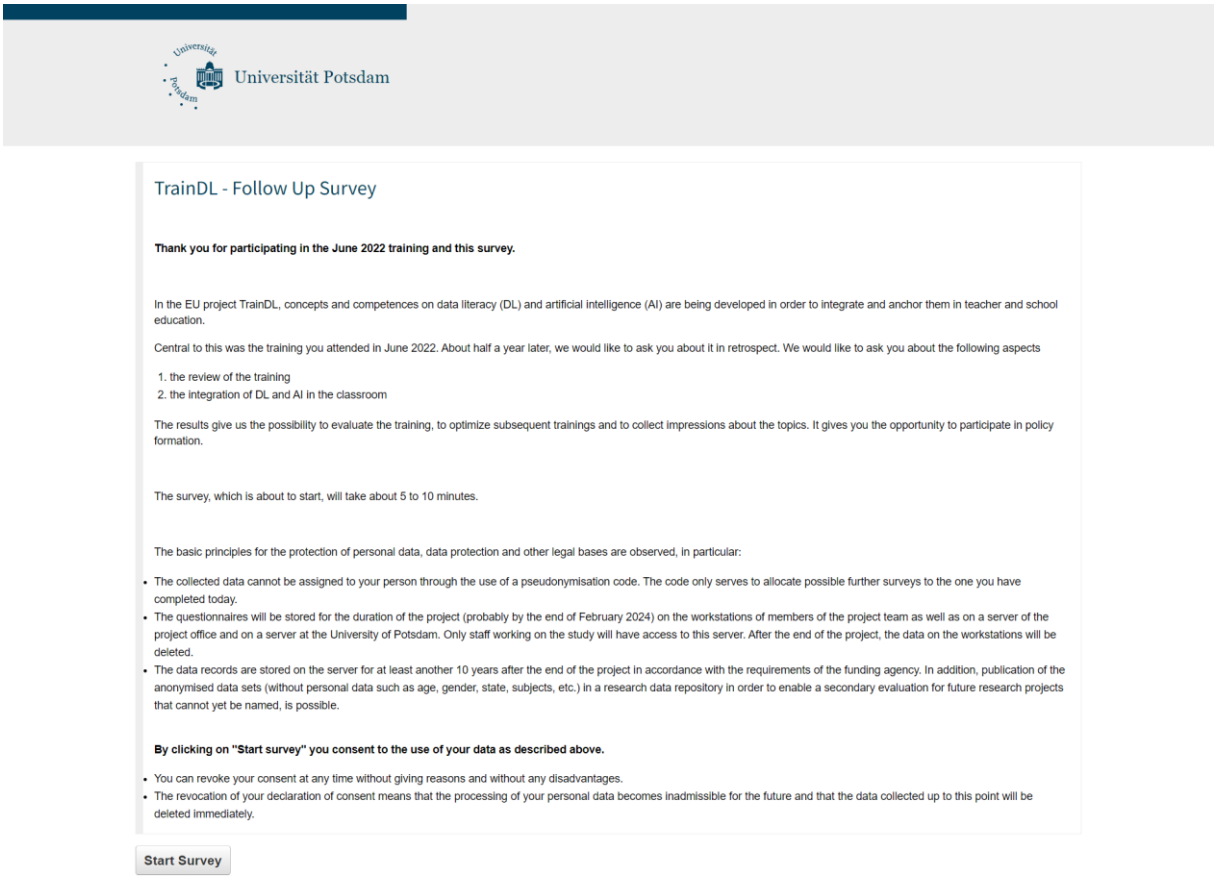
99%

Is there anything else you would like to tell us at the end of the questionnaire/training?

Optional reponse

Back Send questionnaire

10.3 CS training, Berlin (13.06.2022): Translated Version of the follow-up evaluation questionnaire



Universität Potsdam

TrainDL - Follow Up Survey

Thank you for participating in the June 2022 training and this survey.

In the EU project TrainDL, concepts and competences on data literacy (DL) and artificial intelligence (AI) are being developed in order to integrate and anchor them in teacher and school education.

Central to this was the training you attended in June 2022. About half a year later, we would like to ask you about it in retrospect. We would like to ask you about the following aspects

1. the review of the training
2. the integration of DL and AI in the classroom

The results give us the possibility to evaluate the training, to optimize subsequent trainings and to collect impressions about the topics. It gives you the opportunity to participate in policy formation.

The survey, which is about to start, will take about 5 to 10 minutes.

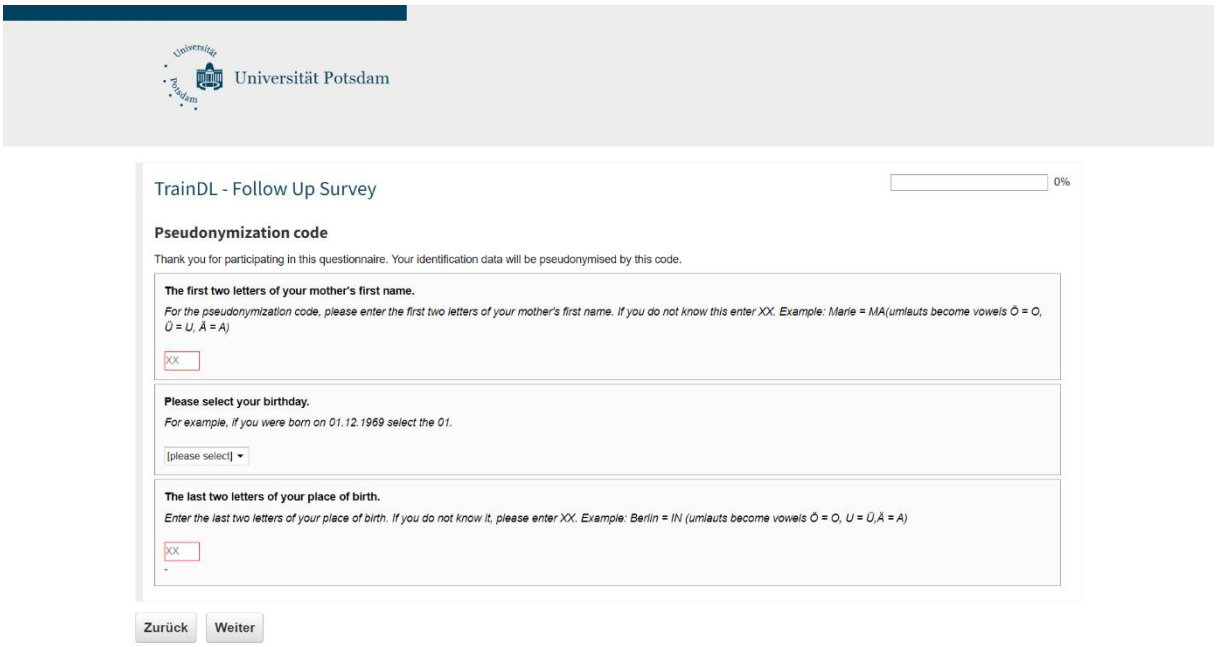
The basic principles for the protection of personal data, data protection and other legal bases are observed, in particular:

- The collected data cannot be assigned to your person through the use of a pseudonymisation code. The code only serves to allocate possible further surveys to the one you have completed today.
- The questionnaires will be stored for the duration of the project (probably by the end of February 2024) on the workstations of members of the project team as well as on a server of the project office and on a server at the University of Potsdam. Only staff working on the study will have access to this server. After the end of the project, the data on the workstations will be deleted.
- The data records are stored on the server for at least another 10 years after the end of the project in accordance with the requirements of the funding agency. In addition, publication of the anonymised data sets (without personal data such as age, gender, state, subjects, etc.) in a research data repository in order to enable a secondary evaluation for future research projects that cannot yet be named, is possible.

By clicking on "Start survey" you consent to the use of your data as described above.

- You can revoke your consent at any time without giving reasons and without any disadvantages.
- The revocation of your declaration of consent means that the processing of your personal data becomes inadmissible for the future and that the data collected up to this point will be deleted immediately.

Start Survey



Universität Potsdam

TrainDL - Follow Up Survey

0%

Pseudonymization code

Thank you for participating in this questionnaire. Your identification data will be pseudonymised by this code.

The first two letters of your mother's first name.

For the pseudonymization code, please enter the first two letters of your mother's first name. If you do not know this enter XX. Example: Marie = MA(umlauts become vowels Ö = O, Ü = U, Ä = A)

XX

Please select your birthday.

For example, If you were born on 01.12.1969 select the 01.

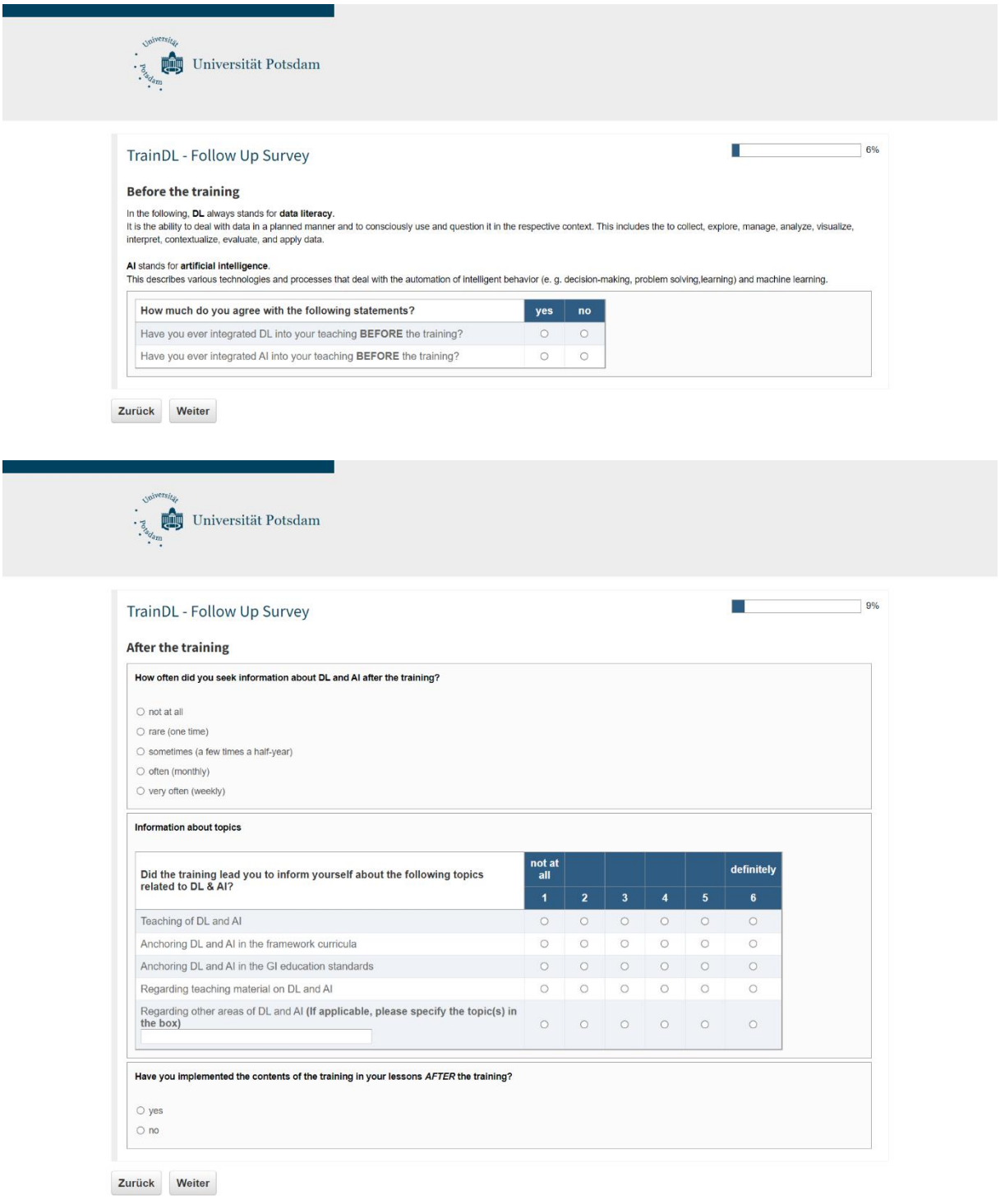
[please select]

The last two letters of your place of birth.

Enter the last two letters of your place of birth. If you do not know it, please enter XX. Example: Berlin = IN (umlauts become vowels Ö = O, U = Ü,Ä = A)

XX

Zurück Weiter



Universität Potsdam

TrainDL - Follow Up Survey

6%

Before the training

In the following, **DL** always stands for **data literacy**.
It is the ability to deal with data in a planned manner and to consciously use and question it in the respective context. This includes the to collect, explore, manage, analyze, visualize, interpret, contextualize, evaluate, and apply data.

AI stands for **artificial intelligence**.
This describes various technologies and processes that deal with the automation of intelligent behavior (e. g. decision-making, problem solving,learning) and machine learning.

How much do you agree with the following statements?	yes	no
Have you ever integrated DL into your teaching **BEFORE** the training?	○	○
Have you ever integrated AI into your teaching **BEFORE** the training?	○	○

Zurück Weiter

Universität Potsdam

TrainDL - Follow Up Survey

9%

After the training

How often did you seek information about DL and AI after the training?

- ○ not at all
- ○ rare (one time)
- ○ sometimes (a few times a half-year)
- ○ often (monthly)
- ○ very often (weekly)

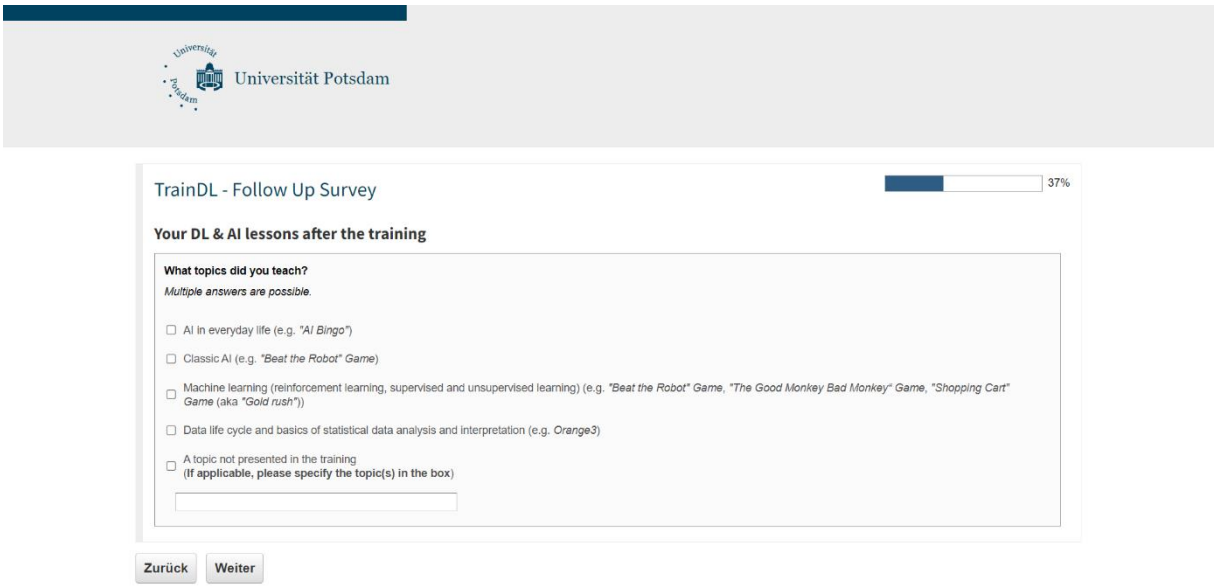
Information about topics

Did the training lead you to inform yourself about the following topics related to DL & AI?	not at all 1	2	3	4	5	definitely 6
Teaching of DL and AI	○	○	○	○	○	○
Anchoring DL and AI in the framework curricula	○	○	○	○	○	○
Anchoring DL and AI in the GI education standards	○	○	○	○	○	○
Regarding teaching material on DL and AI	○	○	○	○	○	○
Regarding other areas of DL and AI **(If applicable, please specify the topic(s) in the box)**	○	○	○	○	○	○

Have you implemented the contents of the training in your lessons *AFTER* the training?

- ○ yes
- ○ no

Zurück Weiter



Universität Potsdam

TrainDL - Follow Up Survey

37%

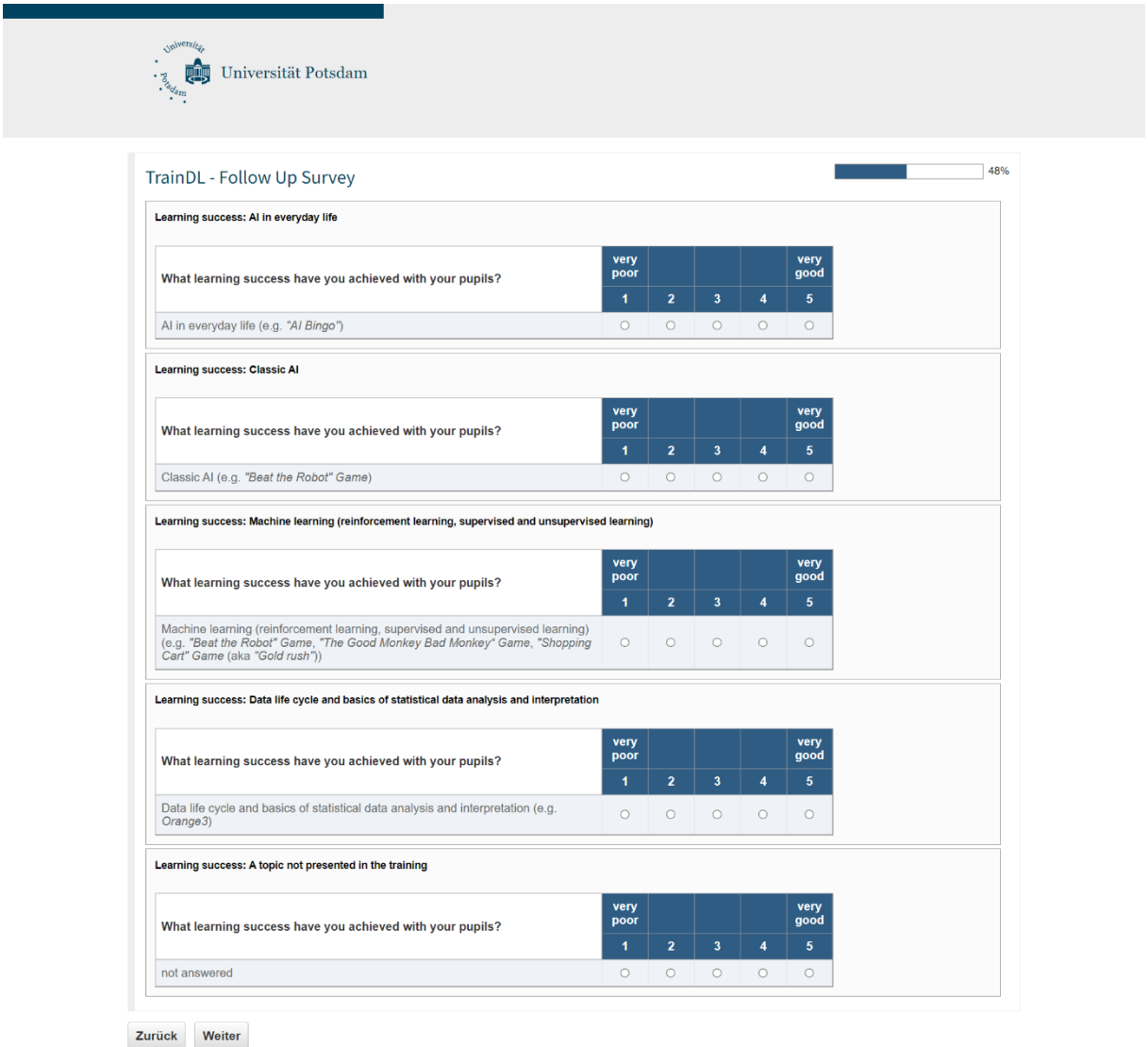
Your DL & AI lessons after the training

What topics did you teach?

Multiple answers are possible.

- ☐ AI in everyday life (e.g. *"AI Bingo"*)
- ☐ Classic AI (e.g. *"Beat the Robot" Game*)
- ☐ Machine learning (reinforcement learning, supervised and unsupervised learning) (e.g. *"Beat the Robot" Game, "The Good Monkey Bad Monkey" Game, "Shopping Cart" Game* (aka *"Gold rush"*))
- ☐ Data life cycle and basics of statistical data analysis and interpretation (e.g. *Orange3*)
- ☐ A topic not presented in the training
  (If applicable, please specify the topic(s) in the box)

Zurück Weiter



Universität Potsdam

TrainDL - Follow Up Survey

48%

Learning success: AI in everyday life

What learning success have you achieved with your pupils?	very poor				very good
	1	2	3	4	5
AI in everyday life (e.g. *"AI Bingo"*)	○	○	○	○	○

Learning success: Classic AI

What learning success have you achieved with your pupils?	very poor				very good
	1	2	3	4	5
Classic AI (e.g. *"Beat the Robot" Game*)	○	○	○	○	○

Learning success: Machine learning (reinforcement learning, supervised and unsupervised learning)

What learning success have you achieved with your pupils?	very poor				very good
	1	2	3	4	5
Machine learning (reinforcement learning, supervised and unsupervised learning) (e.g. *"Beat the Robot" Game, "The Good Monkey Bad Monkey" Game, "Shopping Cart" Game* (aka *"Gold rush"*))	○	○	○	○	○

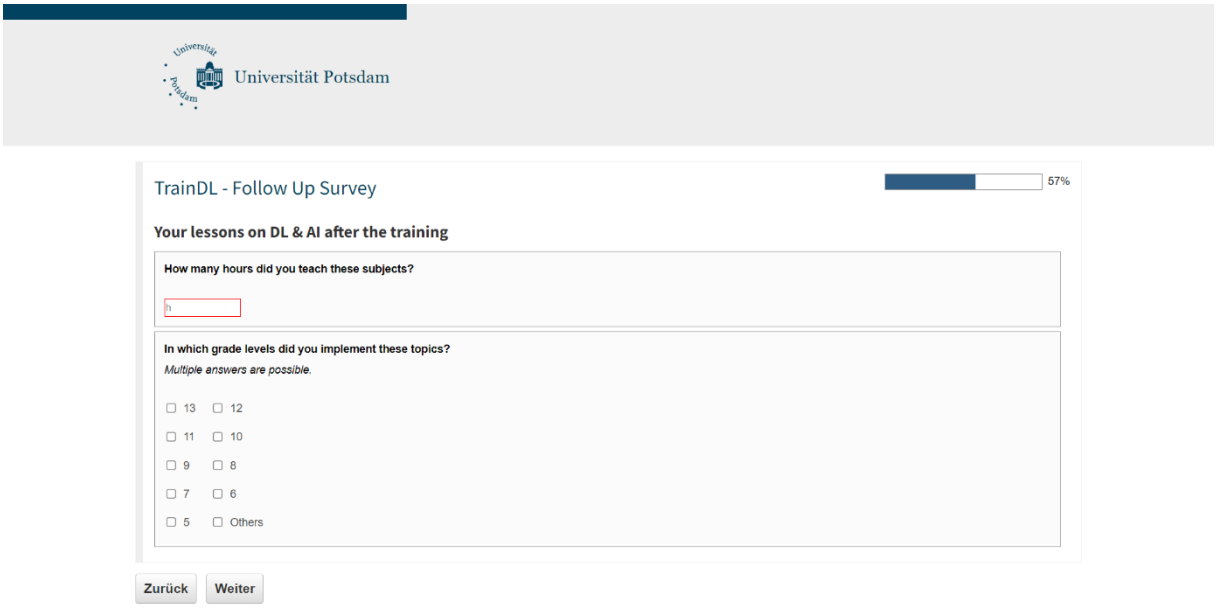
Learning success: Data life cycle and basics of statistical data analysis and interpretation

What learning success have you achieved with your pupils?	very poor				very good
	1	2	3	4	5
Data life cycle and basics of statistical data analysis and interpretation (e.g. *Orange3*)	○	○	○	○	○

Learning success: A topic not presented in the training

What learning success have you achieved with your pupils?	very poor				very good
	1	2	3	4	5
not answered	○	○	○	○	○

Zurück Weiter



Universität Potsdam

TrainDL - Follow Up Survey

57%

Your lessons on DL & AI after the training

How many hours did you teach these subjects?

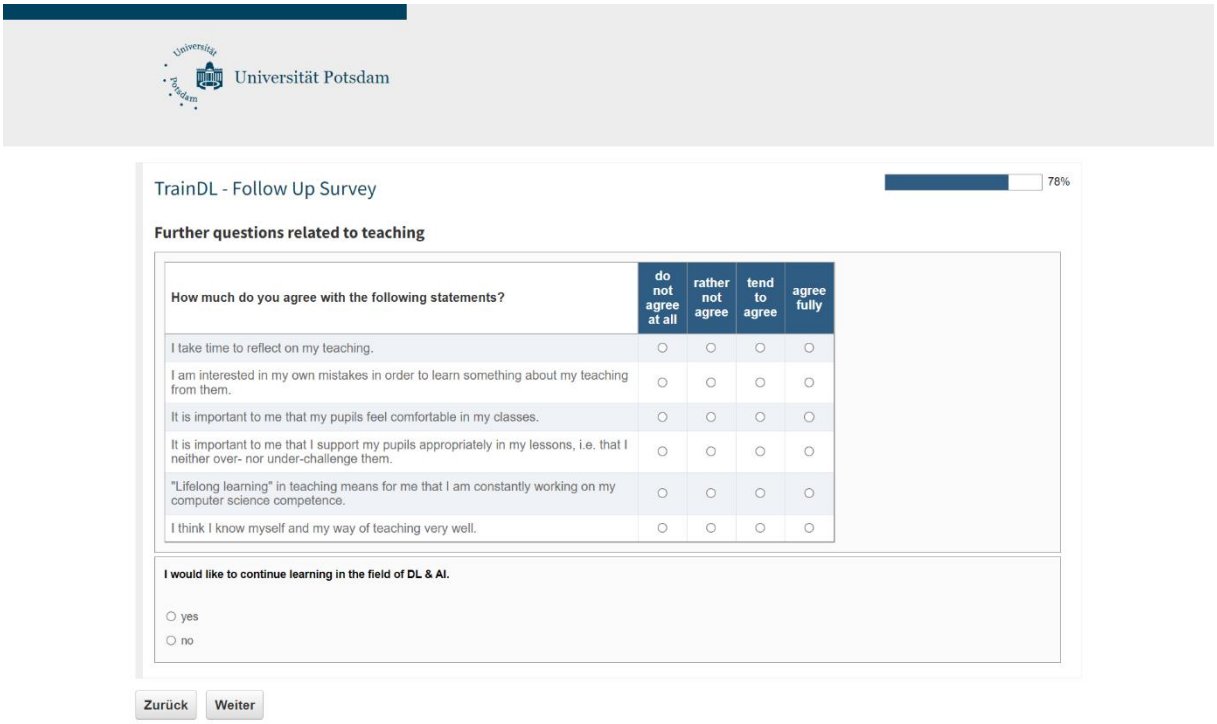
h

In which grade levels did you implement these topics?

Multiple answers are possible.

☐ 13 ☐ 12
☐ 11 ☐ 10
☐ 9 ☐ 8
☐ 7 ☐ 6
☐ 5 ☐ Others

Zurück Weiter



Universität Potsdam

TrainDL - Follow Up Survey

78%

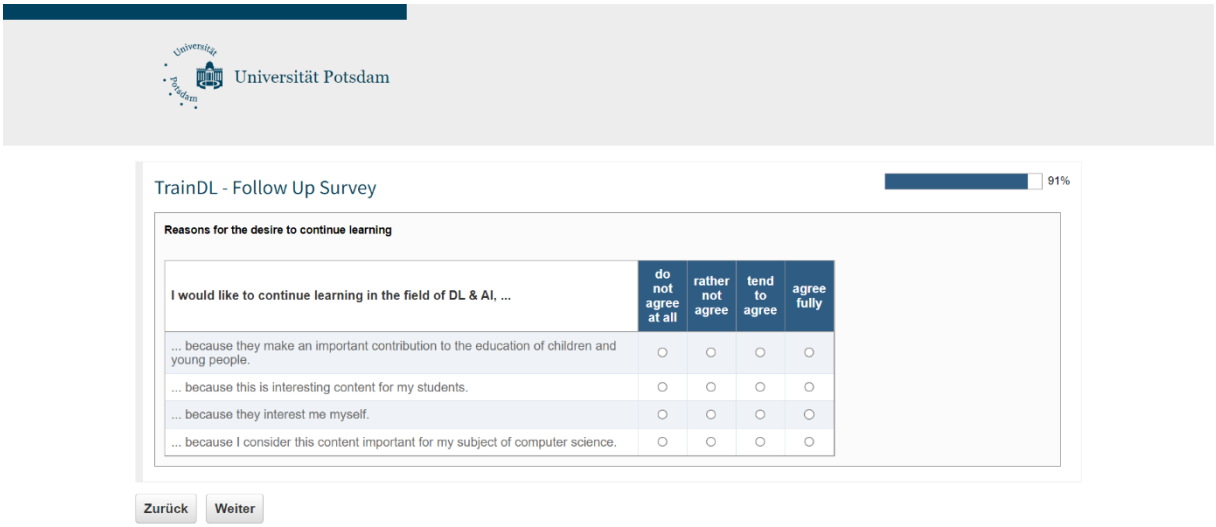
Further questions related to teaching

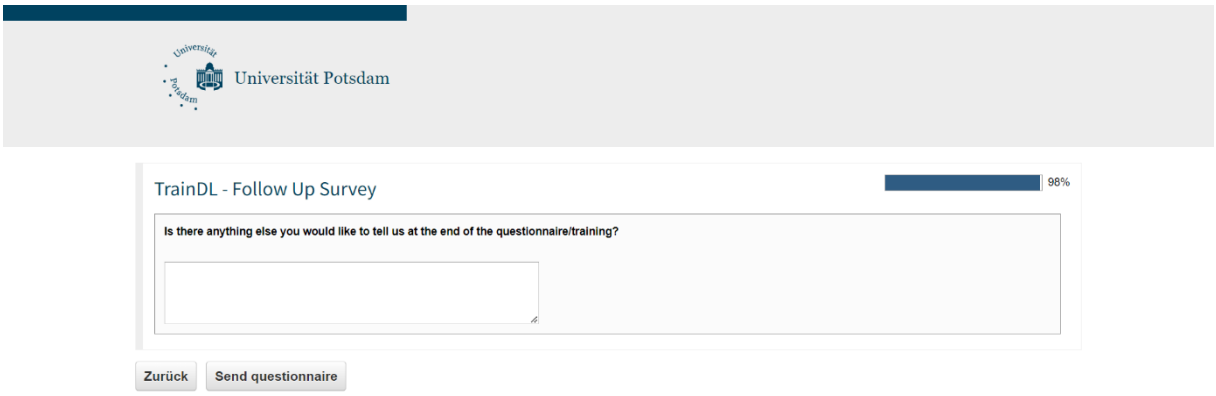
How much do you agree with the following statements?	do not agree at all	rather not agree	tend to agree	agree fully
I take time to reflect on my teaching.	○	○	○	○
I am interested in my own mistakes in order to learn something about my teaching from them.	○	○	○	○
It is important to me that my pupils feel comfortable in my classes.	○	○	○	○
It is important to me that I support my pupils appropriately in my lessons, i.e. that I neither over- nor under-challenge them.	○	○	○	○
"Lifelong learning" in teaching means for me that I am constantly working on my computer science competence.	○	○	○	○
I think I know myself and my way of teaching very well.	○	○	○	○

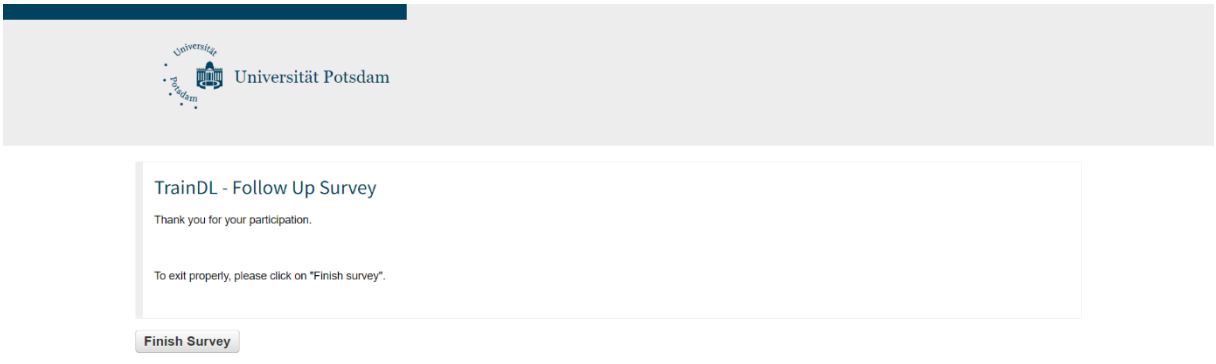
I would like to continue learning in the field of DL & AI.

○ yes
○ no

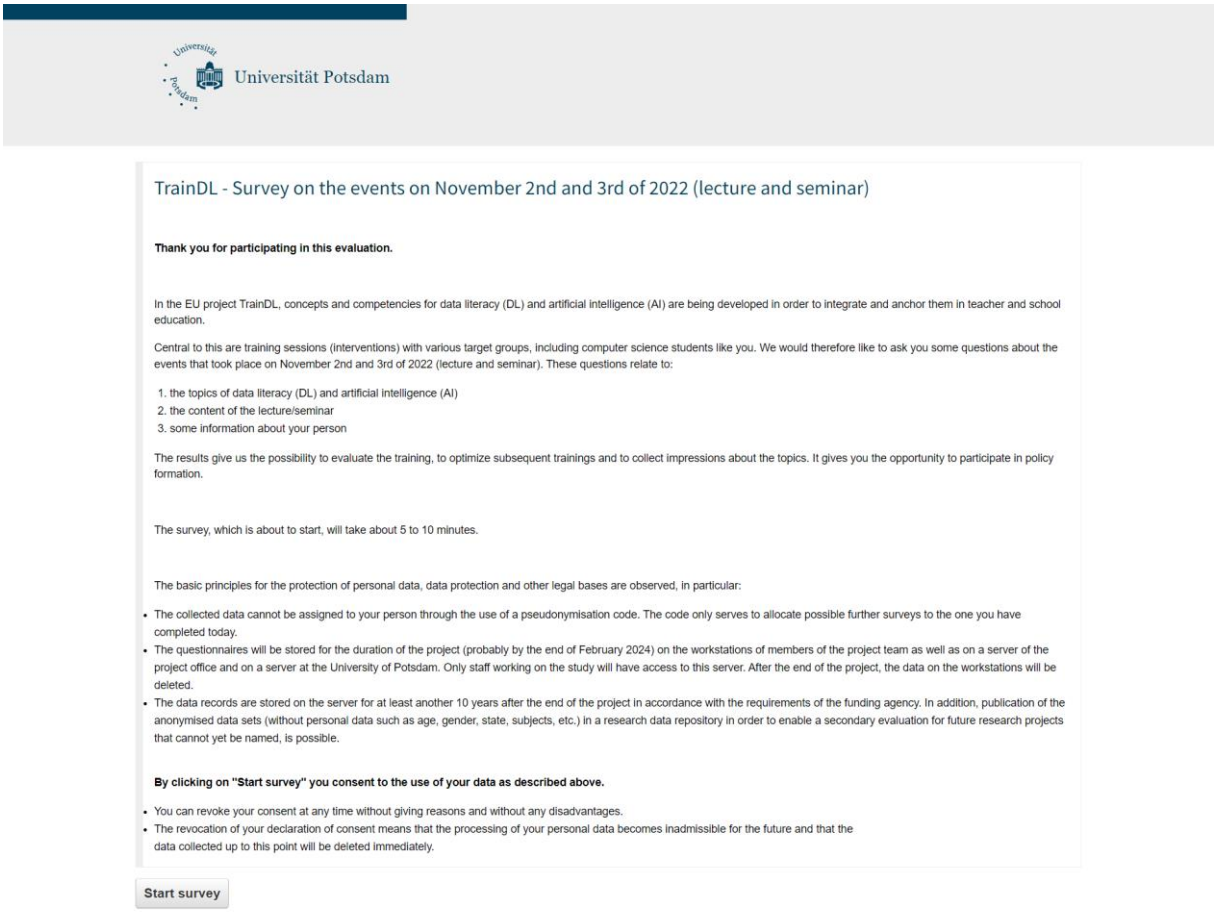
Zurück Weiter


Universität Potsdam
TrainDL - Follow Up Survey
91%
Reasons for the desire to continue learning
I would like to continue learning in the field of DL & AI, ...
do not agree at all
rather not agree
tend to agree
agree fully
... because they make an important contribution to the education of children and young people.
... because this is interesting content for my students.
... because they interest me myself.
... because I consider this content important for my subject of computer science.
Zurück
Weiter


Universität Potsdam
TrainDL - Follow Up Survey
98%
Is there anything else you would like to tell us at the end of the questionnaire/training?
Zurück
Send questionnaire


Universität Potsdam
TrainDL - Follow Up Survey
Thank you for your participation.
To exit properly, please click on "Finish survey".
Finish Survey

10.4 Pre-service CS training, Berlin (11.2022): Translated Version of the post-training evaluation questionnaire



Universität Potsdam

TrainDL - Survey on the events on November 2nd and 3rd of 2022 (lecture and seminar)

Thank you for participating in this evaluation.

In the EU project TrainDL, concepts and competencies for data literacy (DL) and artificial intelligence (AI) are being developed in order to integrate and anchor them in teacher and school education.

Central to this are training sessions (interventions) with various target groups, including computer science students like you. We would therefore like to ask you some questions about the events that took place on November 2nd and 3rd of 2022 (lecture and seminar). These questions relate to:

1. the topics of data literacy (DL) and artificial intelligence (AI)
2. the content of the lecture/seminar
3. some information about your person

The results give us the possibility to evaluate the training, to optimize subsequent trainings and to collect impressions about the topics. It gives you the opportunity to participate in policy formation.

The survey, which is about to start, will take about 5 to 10 minutes.

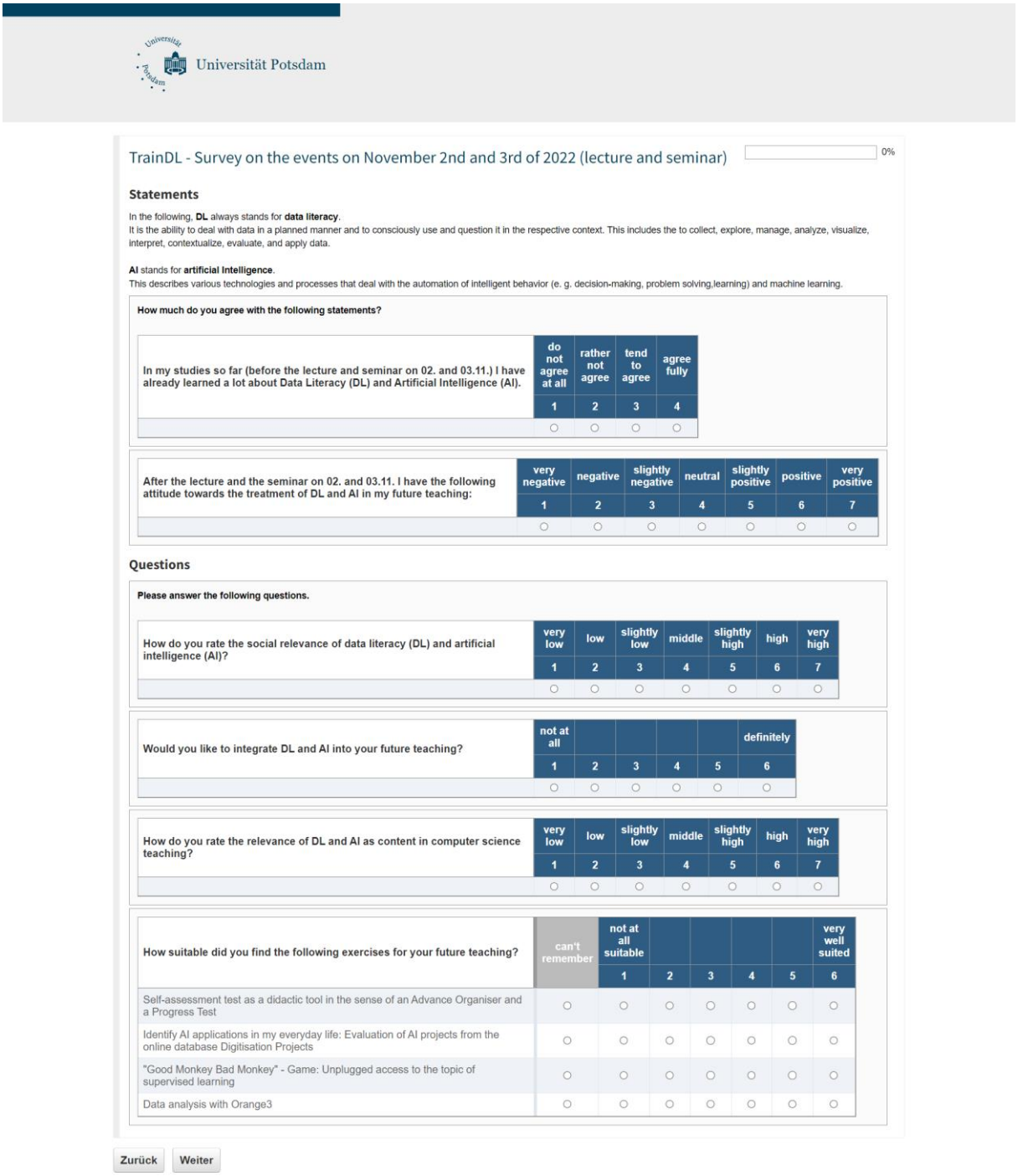
The basic principles for the protection of personal data, data protection and other legal bases are observed, in particular:

- The collected data cannot be assigned to your person through the use of a pseudonymisation code. The code only serves to allocate possible further surveys to the one you have completed today.
- The questionnaires will be stored for the duration of the project (probably by the end of February 2024) on the workstations of members of the project team as well as on a server of the project office and on a server at the University of Potsdam. Only staff working on the study will have access to this server. After the end of the project, the data on the workstations will be deleted.
- The data records are stored on the server for at least another 10 years after the end of the project in accordance with the requirements of the funding agency. In addition, publication of the anonymised data sets (without personal data such as age, gender, state, subjects, etc.) in a research data repository in order to enable a secondary evaluation for future research projects that cannot yet be named, is possible.

By clicking on "Start survey" you consent to the use of your data as described above.

- You can revoke your consent at any time without giving reasons and without any disadvantages.
- The revocation of your declaration of consent means that the processing of your personal data becomes inadmissible for the future and that the data collected up to this point will be deleted immediately.

Start survey



Universität Potsdam

TrainDL - Survey on the events on November 2nd and 3rd of 2022 (lecture and seminar) 0%

Statements

In the following, **DL** always stands for **data literacy**.
It is the ability to deal with data in a planned manner and to consciously use and question it in the respective context. This includes the to collect, explore, manage, analyze, visualize, interpret, contextualize, evaluate, and apply data.

AI stands for **artificial Intelligence**.
This describes various technologies and processes that deal with the automation of intelligent behavior (e. g. decision-making, problem solving,learning) and machine learning.

How much do you agree with the following statements?

	do not agree at all	rather not agree	tend to agree	agree fully
	1	2	3	4
In my studies so far (before the lecture and seminar on 02. and 03.11.) I have already learned a lot about Data Literacy (DL) and Artificial Intelligence (AI).	○	○	○	○

	very negative	negative	slightly negative	neutral	slightly positive	positive	very positive
	1	2	3	4	5	6	7
After the lecture and the seminar on 02. and 03.11. I have the following attitude towards the treatment of DL and AI in my future teaching:	○	○	○	○	○	○	○

Questions

Please answer the following questions.

	very low	low	slightly low	middle	slightly high	high	very high
	1	2	3	4	5	6	7
How do you rate the social relevance of data literacy (DL) and artificial intelligence (AI)?	○	○	○	○	○	○	○

	not at all					definitely
	1	2	3	4	5	6
Would you like to integrate DL and AI into your future teaching?	○	○	○	○	○	○

	very low	low	slightly low	middle	slightly high	high	very high
	1	2	3	4	5	6	7
How do you rate the relevance of DL and AI as content in computer science teaching?	○	○	○	○	○	○	○

How suitable did you find the following exercises for your future teaching?	can't remember	not at all suitable					very well suited
		1	2	3	4	5	6
Self-assessment test as a didactic tool in the sense of an Advance Organiser and a Progress Test	○	○	○	○	○	○	○
Identify AI applications in my everyday life: Evaluation of AI projects from the online database Digitisation Projects	○	○	○	○	○	○	○
"Good Monkey Bad Monkey" - Game: Unplugged access to the topic of supervised learning	○	○	○	○	○	○	○
Data analysis with Orange3	○	○	○	○	○	○	○

Zurück Weiter

10.5 CS in-service training, Berlin (13.06.2022): The post-training interview guide

Interview guide

You just attended a teacher training on Data Literacy (DL) and Artificial Intelligence (AI):

1. How did it come about?

If not answered, address the following (follow-up questions):

1.1 What were your expectations of the training?

1.2 Was the training appropriately designed according to your previous experience with AI & DL?

yes	no
1.2.1 Describe what you liked about it.	1.2.2 Describe what you did not like about it.

1.3 After the training, how would you personally rate the difficulty of the topic for you?

Transition - Previous practical experience with topics in continuing education:

2. Have you ever integrated content about AI and/or DL into your teaching?

yes	no
2.1 How was your experience? What was good? Were there any problems? *If not answered:* 2.2 Which topics from the fields of DL or AI did you cover? 2.3 After the training, do you feel able to integrate DL and AI more effectively in the classroom?	2.4 What keeps you from introducing DL & AI in the classroom? *If not answered:* 2.5 After this training, how ready do you feel to integrate DL & AI in the classroom?

Transition - before teaching side viewed; now other level:

3. As a teacher, what do you think about establishing AI and DL in the curricular frameworks?

If not answered:

3.1 What do you think about making AI and DL an integral part of the teacher training?

For questions/no answer: module/event, seminars, aspects, etc.

3.2 What steps would be needed to embed AI & DL topics in the classroom?

For questions /no response: e.g., recommendations, guidelines/directions, curricular framework changes.

3.3 What steps could be taken to integrate the topic of AI and DL into your school's internal curriculum?

For questions /no response: explain internal school curriculum differently, e.g. school action concept (without conveying that respondent would not understand the term)

3.4 To what extent do you anticipate certain institutional obstacles or difficulties?

For questions/no answer: acceptance issues, transition phases, convincing of meaning and purpose would have to be done, etc.

Transition - Consequences of the changes to the curricular frameworks just discussed:

4. What changes might the integration of AI and DL bring at different levels?

If not answered, be sure to address the following (follow-up questions):

4.1 among the students?

4.2 with the school authorities?

4.3 In society at large?

Transition - Back to training/intervention today:

5. How can future training be improved or made more efficient?

If not answered, be sure to address the following (follow-up questions):

5.1 Length

5.2 Contents

If there is still enough time:

5.3 Format

5.4 Interaction (proportion of exercises - frontal teaching?)

Transition - So far both teaching perspective and levels beyond; now students/students again:

6. After the training, how would you rate the communicability of AI and DL to students?

Transition - Coming slowly to the end:

7. If you had one wish for education policy: What would you change about DL & AI in the classroom?

If not answered, be sure to address the following (follow-up questions):

7.1 Are you sure you don't want to add anything?

Transition - to conclusion:

8. Independently, is there anything else you would like to share with us? Do you have any suggestions?

10.6 CS in-service training, Berlin: Follow-up guide for interviews roughly a half year after the initial training on 13.06.2022

Interview guide

You attended in a teacher training on Data Literacy (DL) and Artificial Intelligence (AI) in June of this year.

1. Looking back, what have you taken away from the training for yourself and/or your teaching?

If not answered:

- 1.1 Self
- 1.2 Training
- 1.3 Which **specific aspect** has stuck with you the most?
- 1.4 To what extent have you had the opportunity to further engage with AI and DL after the training?

Transition - Now we come to a potential integration of Data Literacy and Artificial Intelligence in the classroom.

2. Have you integrated AI and DL into your teaching, or do you plan to integrate these topics?

yes	no
2.1 How was your experience? *If not answered:* 2.2 Regarding time frame *For questions/no answer: e.g., double lesson* 2.3 Regarding the performance/grade level *For questions/no answer: e.g., 12th grade, advanced course or basic course* 2.4 Covered topics/content? *For questions/no answer: e.g., Machine Learning in the context of databases* 2.5 Method of knowledge dissemination/teaching? *For questions/no answer: e.g., materials, methods* 2.6 Was material from the training used? *For questions/no answer: e.g., material related to robot monkey chess game ("Beat the robot")*	2.7 Are there **specific reasons** for it? *For questions/no answer: e.g., current school year, time constraints* 2.8 What could **help with the integration**?

Transition - Next, we explore the potential for transformation of DL and AI.

3. Where do you see further potential for change regarding these topics?

If not answered, address the following (follow-up questions):

- 3.1 Past training
- 3.2 In general

For questions/no answer: e.g, society, school, framework curriculum

Transition - Coming slowly to the end:

4. If you had one wish for education policy: What would you change about DL & AI in the classroom?

Transition - to conclusion:

5. Independently, is there anything else you would like to share with us? Do you have any suggestion?