2.9. Report on the Third Round of Interventions

Version: April 2024

https://train-dl.eu

Prof. Dr. Ralf Romeike, Viktoriya Olari, Till Zoppke | Freie Universität Berlin

Summary

This report details the implementation and evaluation of the third round of interventions (all teachers) that was conducted from August 2023 until January 2024. Also included are the results of a weekly seminar for CS teacher students that took place from April to July 2023. In total, 9 interventions with 74 participating CS teachers (including 15 pre-service) and 95 STEAM, languages, humanities and sports teachers were conducted across Germany, Austria, and Lithuania.

The third round of interventions was based on a modular workshop concept so that each intervention could be adapted to the learning objectives, the target group and the time frame available. As for the previous rounds, we outline the workshop concepts and introduce newly developed activities. Among other things, ethical aspects (algorithmic bias), computer vision, linear regression, AI agents and knowledge bases were comprehensively updated. In addition, there are new activities for dealing with AI chatbots and developing teaching ideas for their use.

The workshops were evaluated by the University of Potsdam (WP4) with a survey. We also conducted interviews with the individual participants or with multiple participants at the same time in the form of focus group sessions. The evaluation results can be found in the respective deliverables of WP4. The report describes how the interventions were implemented, including facilitators reflection on them and feedback of the participants encountered within the workshops.

Table of Contents

Introduction

Recent breakthroughs in generative AI and the widespread application of systems based on it make their societal relevance all the more evident. To enable students to follow this development, AI and data literacy should be integrated into the school curriculum in its technical, application-oriented and ethical dimensions. One key to the successful integration of AI and data literacy, however, is the effective preparation of teachers. Therefore, teachers of all subjects should be empowered to guide their students.

In the TrainDL project, we design, implement and evaluate teacher trainings as a means of concluding policy guidelines for stakeholders. The trainings are practice-oriented and focus on activities that can be used in the classroom. Based on the Dagstuhl Triangle Framework teachers acquire technological, application-oriented and ethical content knowledge as well as pedagogical knowledge on how to integrate the learned content into the classroom.

In the first round of interventions, reported in Deliverable D2.3, we conducted trainings with secondary CS teachers. In the second round of interventions, reported in D2.6., we continued to target CS teachers while expanding the scope to include secondary STEAM teachers. In the third intervention round, on which we report here, we targeted teachers from all subjects (CS, STEAM, languages, humanities, sports). In the third round, we implemented and evaluated the following seven interventions:

23.08.2023 | Düsseldorf | Secondary school teacher-trainers
05.09.2023 | Berlin | STEAM teachers
18.09.2023 | Berlin | CS teachers
23.10.2023 | Austria | Secondary school teachers
24.11. - 26.11.2023 | Zeitz | CS teachers
09.01.2024 | Vilnius | Secondary school teachers
26.01.2024 - 28.01.2024 | Meißen | Secondary school teachers

In addition, besides the regular interventions in the third round, to pilot our teaching materials and to better understand teachers' opinions about our instructional approach and design, we conducted two additional trainings with **in- and pre-service teachers** from secondary schools on the following dates:

25.04. - 18.07.2023 | Berlin | CS Pre-Service Teachers
18.10.2023 | Berlin | STEAM teachers

In what follows, we first briefly present the **workshop format** for the interventions and the process we followed to **evaluate the effectiveness of the sessions**. We outline the number of teachers who participated in the trainings and in the evaluation process. We then present the implementation details for each training session, along with **feedback**

from the facilitators who had conducted the sessions. Finally, we summarize the challenges encountered during the third round of interventions.

Design of the Training Sessions

In this chapter, the design of the training sessions will be described. It begins with a description of the activities and learning materials that were newly developed in the third round of interventions. Next, we introduce the training formats and present how the effectiveness of the training sessions was measured.

Activities and Learning materials

In the third round of interventions, 12 activities from the prior rounds of interventions were updated or newly developed. A complete overview of all activities from the three intervention rounds can be found in the Final Experience Report (D2.11). In the following, we give an overview of the activities for the third round of interventions.

AI Quiz cards

The activity "AI Quiz", which we used in our interventions in the first round, was further developed into a card game in the third round. It consists of 10 cards, each with an AI system depicted on the front side and a brief description of the system on the back. The task for the learners is to sort the cards according to how much AI they contain.

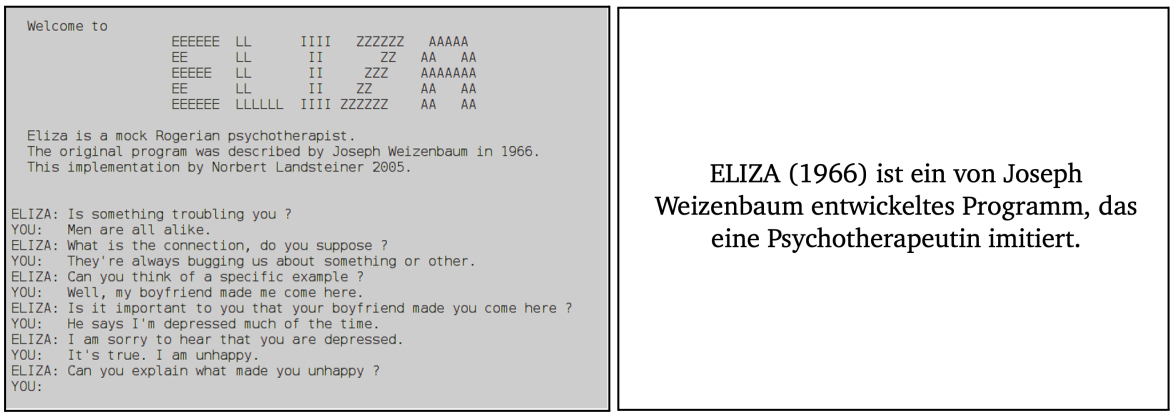


Figure 1: Front and back of the AI quiz card for Joseph Weizenbaum's ELIZA

A learning objective is for the participants to think about what AI is and what defines an AI system. They discuss this in groups of 3-4 people while sorting the cards. In an alternative variation, they are instead asked to indicate for each system whether it is AI (yes / no / maybe) and write a sentence to explain their reasoning.

For resolution and consolidation, there is a slide set that specifies relevant features and AI concepts for each quiz card, such as Weizenbaum's ELIZA and the “ELIZA effect”, which states that people consider a system to be more intelligent than it actually is, but this attribution fades when they have insight into how the system is implemented. The overarching learning goals of the activity are to provide an overview of the different types of AI systems and to achieve a common understanding of what AI is.

Statistics Unplugged

As an ice-breaker and for getting to know each other, an activity was used in which the participants positioned themselves in the room in response to a question. This is also known as a "human barometer". In our context of a training on data literacy, we call it "Statistics Unplugged" because data is being collected by the participants' positioning. In the further course of the workshop, reference can be made to this activity, for example as an example of the difference between categorical or numerical data.

AI and Music (Playful showcases of creative AI)

We have developed this activity specifically for a training attended by music teachers. It is a self-directed activity guided by a worksheet, where learners use their digital devices (PCs, tablets or smartphones are possible) to try out three musical AI applications. An AI system takes on the role of a musician in a musical dialogue, composing a four-part chorale based on a melody given by the user, or improvising harmonies. These tools demonstrate the possibilities of AI systems in the field of music.

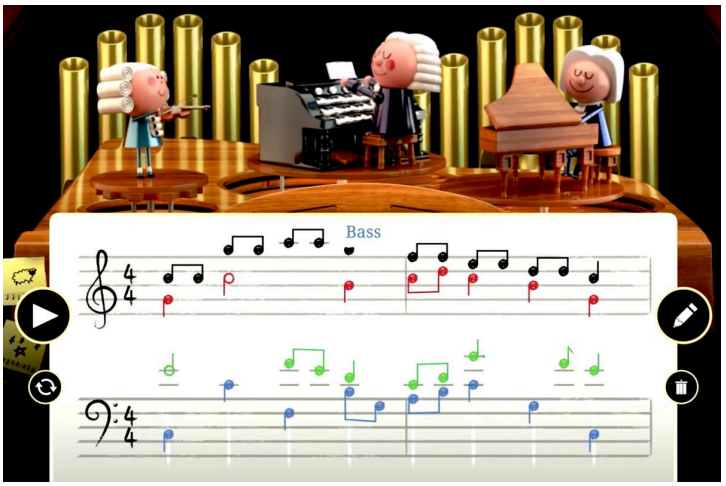


Figure 2: Screenshot of the Johann Sebastian Bach Doodle by Google[1], one of three showcases that are part of the “AI and Music” activity

Material recognition (Image Classification with Teachable Machine)

In this tangible activity students develop a computer vision system based on Teachable Machine[2] system. Implementing a recycling robot, the system should be able to differentiate between plastic and other materials. To do this, students are provided with bags containing examples (plastic bottles, rocks, pieces of wood, leaves). Guided by a worksheet, they use the camera on their PCs to take pictures and train corresponding classes for image recognition. The system will then be tested with additional objects and its accuracy is calculated.

[1] https://doodles.google/doodle/celebrating-johann-sebastian-bach/
[2] https://teachablemachine.withgoogle.com/

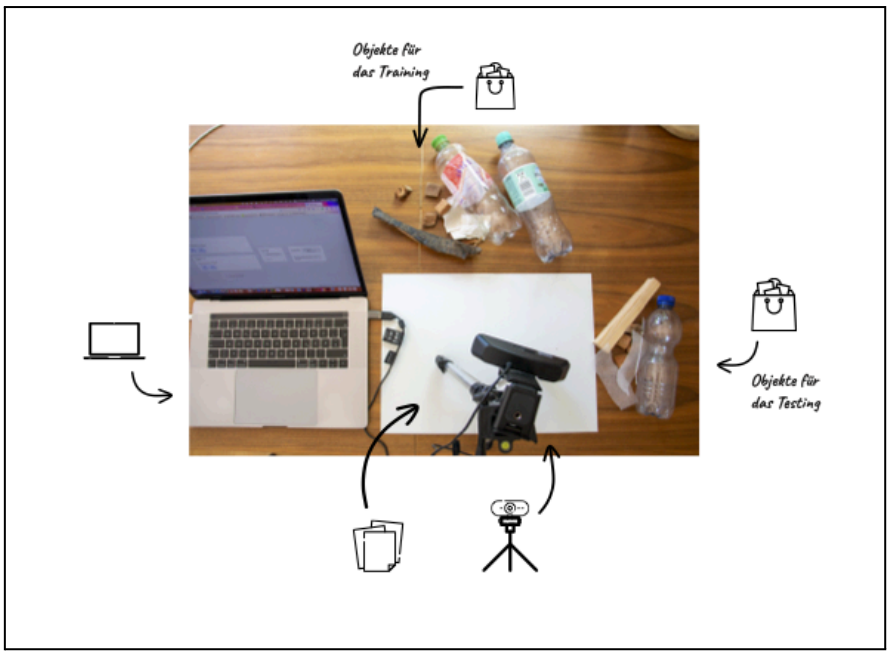


Figure 3: Schematic work setup for the activity Material control

Malaria (Image Classification with Teachable Machine)

This activity is a variation of the activity "Material Recognition" introduced above. It takes into account the feedback of the participants from one of the interventions that not every classroom is equipped with a camera-enabled PC. Accordingly, image recognition is trained on a dataset of prepared images. Here, a dataset with microscopic images of red blood cells is available, some of which are infected with malaria pathogens. Alternative datasets include images of diseased and healthy tomato plants, images of bees with differently filled pollen baskets, and pictures of historical coins. The process of the activity is identical to material recognition, with the data acquisition (creation of image materials) being omitted here.

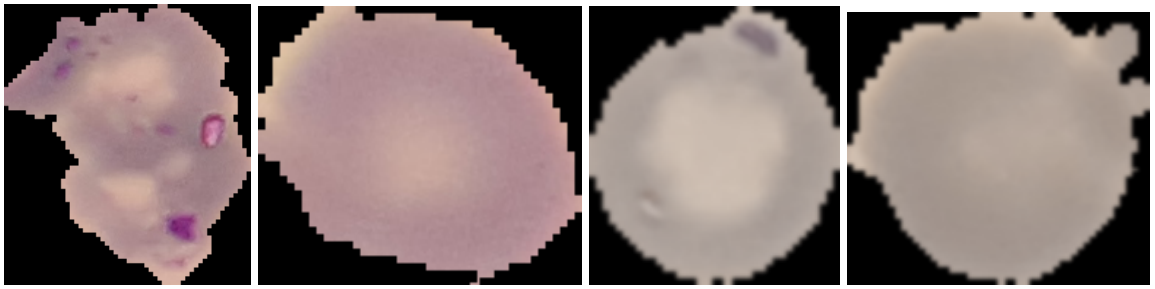


Figure 4: sample data for the activity “Malaria”, including images of healthy and of parasitized blood cells

Food quality control

In this activity, students can learn about three variations of algorithmic bias. They are first given the task of training image recognition that distinguishes ripe bananas from overripe or spoiled ones. For image recognition, a web application[3] of the University of Eastern

[3] https://tm.gen-ai.fi/about

Finland is used. Unlike Google's Teachable Machine, it also runs on iPads and smartphones, making it flexible to use.

The students train the image classifier, and during testing they find a high error rate. Guided by the worksheet, they look at the training data and gradually discover three types of bias. Selection bias occurs when some groups are over- or underrepresented in a training dataset. Framing bias occurs when the images in the dataset only contain one view, scene, or perspective during capture or editing and can therefore convey different messages. Labeling bias occurs when the labeling process during the creation of the training dataset introduces a bias and the data labels deviate from the ground truth.

The system must therefore be improved and retrained. Such loops are typical in the development of machine learning systems. One learning objective is also that for faulty ML systems, data errors are often the cause. The exercise is rounded off with an overview of the possible applications for AI systems along the food chain (agriculture, food industry).

Figure 5: Title page of the worksheet for the activity “Food control” on Algorithmic Bias

Abalone (Linear Regression with Orange3)

For designing AI and data activities in non-computer science school subjects, it is a challenge to embed them into the respective subject curriculum. We researched suitable datasets for natural science and picked a dataset of sea snails. The goal of the activity is to

train an AI model that predicts the age of a sea snail based on six variables such as size, length, or weight.

Figure 6: Title page of the worksheet “How old is the Abalone?” on multiple linear regression in orange3

For this purpose, students create a data workflow in Orange3 and use linear regression to predict the age of a sea snail. An important learning objective is to experience data analysis as a tool for knowledge generation. Conceptually, linear regression builds on linear functions. The machine learning method is visually well representable, making it suitable to be taught by non-computer science teachers as well, especially by math teachers.

In biology class, the activity can be integrated into a lesson series on marine life, and the teacher can also introduce the sea snail as a living organism. For other subjects, the activity may be transferred to other domains and use other datasets.

SuperCodingBall (AI Agents)

In the probably most famous monograph on Artificial Intelligence,[4] is the paradigm of "intelligent agent" is central to the definition of AI. In order to bring this concept into the classroom, we have developed an activity based on the open-source football simulation

[4] Russel, Norvig: Artificial Intelligence: A Modern Approach. https://aima.cs.berkeley.edu/=

SuperCodingBall.[5] In this simulation, teams of four players compete against each other, whose behavior is controlled by intelligent agents. The task is to implement these agents in the programming language Scratch. The students are guided through a worksheet.

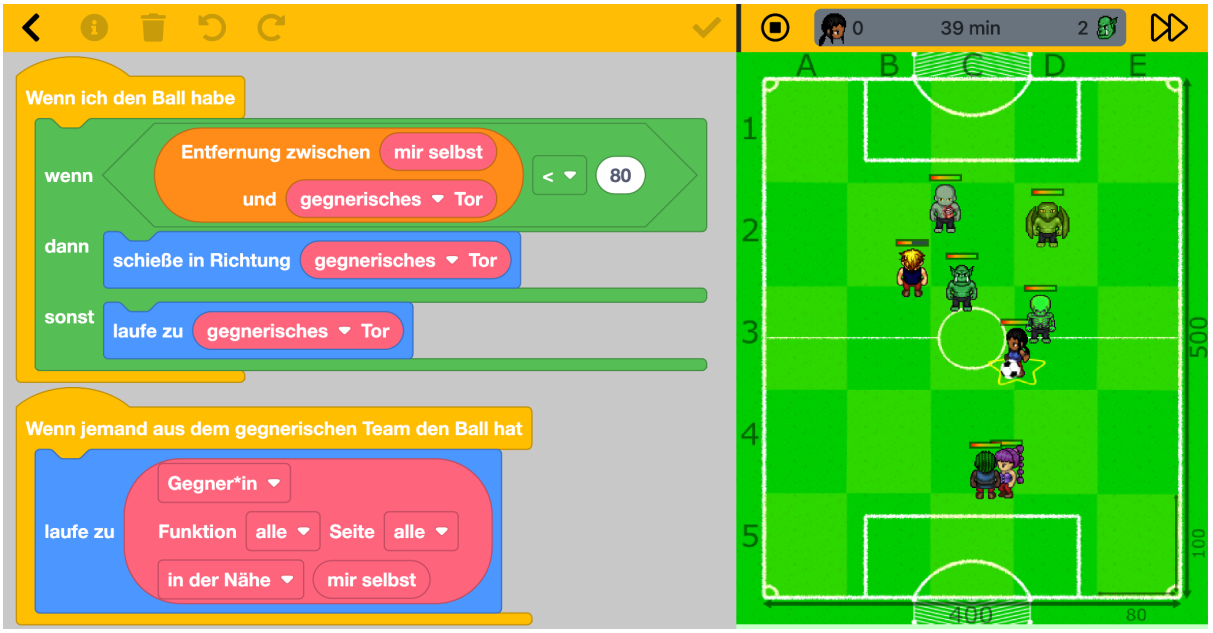
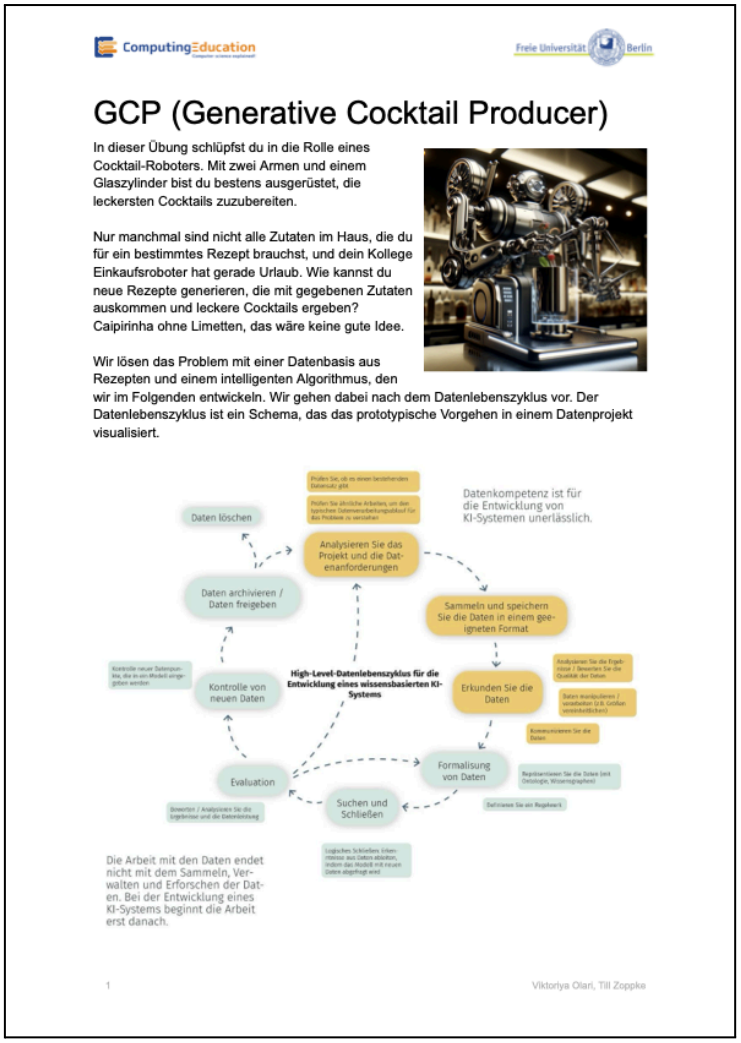


Figure 7: Screenshot of a simple intelligent agent playing simulated soccer in SuperCodingBall

Generative Cocktail Producer (Knowledge base)

In order to cover the field of symbolic AI, we designed an activity in which students can create non-alcoholic cocktail recipes using a knowledge base. The process includes all phases according to the data lifecycle. In the data acquisition phase, students collect recipes (ingredients and quantities) in a table. They then transfer this table to a knowledge graph, where all ingredients are connected to each other that occur in a recipe. The more common the occurrence, the higher the edge weight. From this knowledge base, students can then create new recipes, for example by starting with any ingredient and following the edge with the strongest weight four times.

[5] https://www.supercodingball.com/


ComputingEducation

Freie Universität Berlin

GCP (Generative Cocktail Producer)

In dieser Übung schlüpfst du in die Rolle eines Cocktail-Roboters. Mit zwei Armen und einem Glaszylinder bist du bestens ausgerüstet, die leckersten Cocktails zuzubereiten.

Nur manchmal sind nicht alle Zutaten im Haus, die du für ein bestimmtes Rezept brauchst, und dein Kollege Einkaufsroboter hat gerade Urlaub. Wie kannst du neue Rezepte generieren, die mit gegebenen Zutaten auskommen und leckere Cocktails ergeben? Caipirinha ohne Limetten, das wäre keine gute Idee.

Wir lösen das Problem mit einer Datenbasis aus Rezepten und einem intelligenten Algorithmus, den wir im Folgenden entwickeln. Wir gehen dabei nach dem Datenlebenszyklus vor. Der Datenlebenszyklus ist ein Schema, das das prototypische Vorgehen in einem Datenprojekt visualisiert.

Daten löschen
Analysieren Sie das Projekt und die Datenanforderungen
Daten archivieren / Daten freigeben
Sammeln und speichern Sie die Daten in einem geeigneten Format
Kontrolle von neuen Daten
High-Level-Datenlebenszyklus für die Entwicklung eines wissensbasierten KI-Systems
Erkunden Sie die Daten
Evaluation
Formalisung von Daten
Suchen und Schließen
Datenkompetenz ist für die Entwicklung von KI-Systemen unerlässlich.
Die Arbeit mit den Daten endet nicht mit dem Sammeln, Verwalten und Erforschen der Daten. Bei der Entwicklung eines KI-Systems beginnt die Arbeit erst danach.
1
Viktoriya Olari, Till Zoppke

Figure 8: Title page of the worksheet for the activity “Generative Cocktail Producer” on knowledge-based AI

On the way to the finished recipe, students will make decisions and solve problems. They must clean the data, as one of the recipes is a Margherita pizza, which should not be mixed into a cocktail. Furthermore, decisions must be made about the ingredients during tabular recording, for example whether "lemon juice" is the same as "juice of a lemon". The units of measurement must also be standardized. Finally, the knowledge base must be expanded with quantities and the algorithm must be adjusted accordingly.

Pizza price prediction

This activity was designed specifically for the weekend intervention in Zeitz and is discussed in the respective paragraph.

Prompt Battle

In this activity[6], two players compete against each other. They each formulate a prompt for an image generator (Midjourney, DALL-E) for a given task. The audience decides which of the

[6] https://promptbattle.com/

generated images is better. The goal of the exercise is to get creative with AI. Additionally, students can learn prompt techniques from each other.

AI-PACK

At our last intervention, a weekend workshop for all teachers in January 2024, we tried out an activity based on the AI competence model AI-PACK.[7] This competence model is based on the DPACK and considers that AI can be both a subject of instruction (learning about AI) and a teaching tool (learning with AI). Within the subject matter, the three perspectives of the Dagstuhl Triangle must be taken into account: the technological, socio-cultural, and application-oriented perspectives.

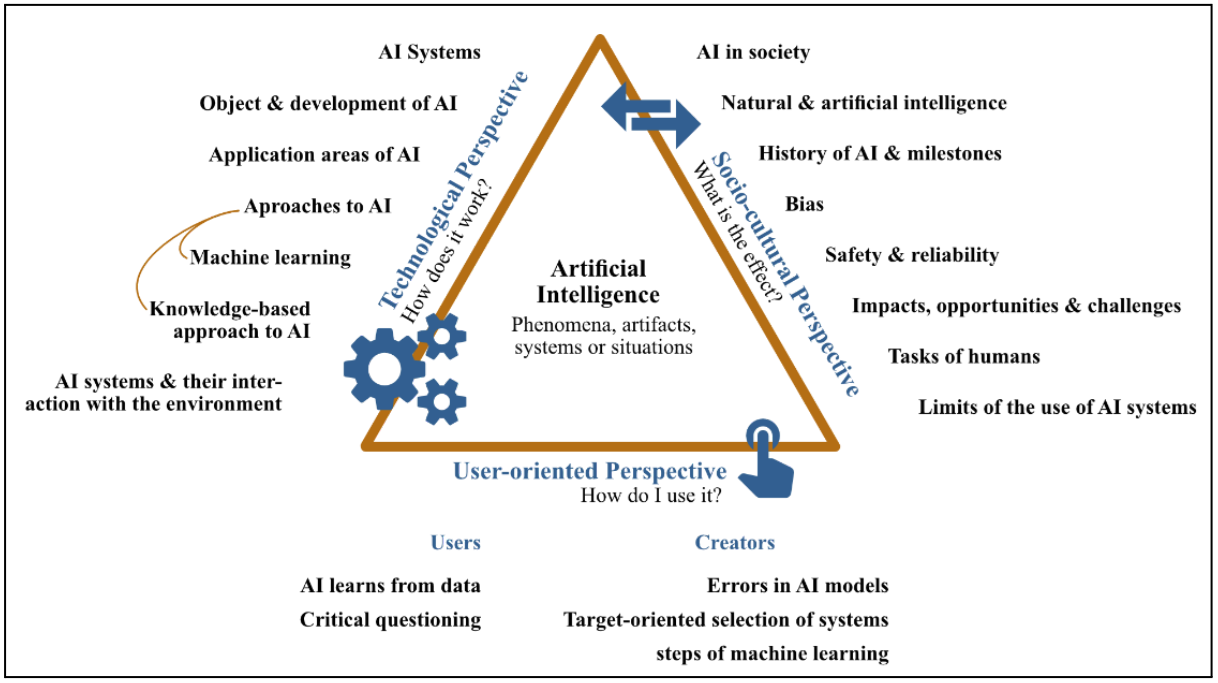


Figure 9. AI competencies in the perspectives of the Dagstuhl triangle.[8]

For each subject, it is now interesting to see which AI phenomena from the students' own experiences can be used in teaching. For example, in politics class, fake news could be addressed, and in history, the photogrammetric recording of ruins. Teachers address this question in group work and then present their results afterwards.

Expert sessions

As in the first weekend workshop (Heidelberg 2022), we invited experts to our weekend workshops in the third round of interventions. The experts presented their respective research to the participants and demonstrated how specialized and diverse the field of AI

[7] https://dl.gi.de/items/6bb42c88-f06b-4b9b-accd-47da7b8d7ed6

[8] Michaeli, T., Romeike, R., Seegerer, S. (2023). What Students Can Learn About Artificial Intelligence – Recommendations for K-12 Computing Education. In: Keane, T., Lewin, C., Brinda, T., Bottino, R. (eds) Towards a Collaborative Society Through Creative Learning. WCCE 2022. IFIP Advances in Information and Communication Technology, vol 685. Springer, Cham. https://doi.org/10.1007/978-3-031-43393-1_19

is. We also held a panel with two experts at one of the workshops. Participants asked any questions about AI and data literacy that interested them.

AI Chatbots and their use in education

STEAM teachers who do not teach computer science have different requirements for training in AI and data literacy. It is important for them to use data and AI systems competently in their professional lives. This is especially important in the classroom, to teach students how they can use those systems competently.

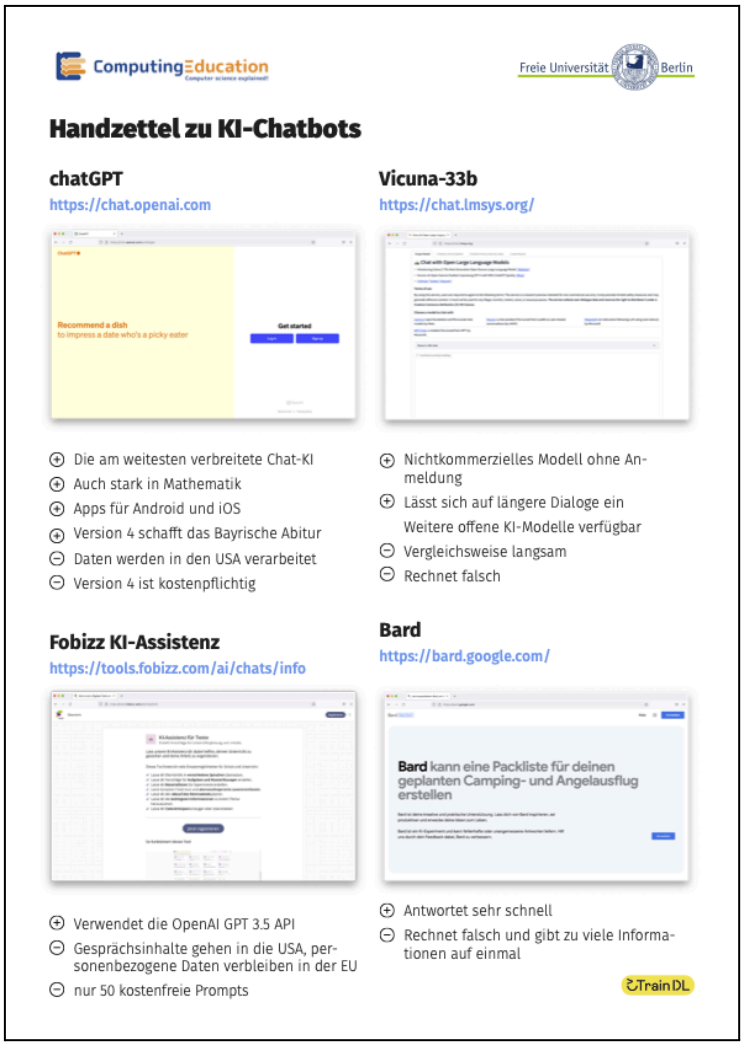


Figure 10: Handout on AI chatbots

For these teachers, upon their explicit request, we have developed an activity on AI chatbots. It starts with the teachers exchanging with each other how they use chat AIs professionally (in class, but also for preparation or professional communication). The teachers form groups based on their subjects, or are assigned so that in each group there are experienced and inexperienced teachers together.

The groups then share their use cases in the plenum. In a subsequent presentation, it is explained what an AI chatbot is and what needs to be considered when using it, including further examples from practical use in schools. Led by the participants' questions, myths about AI chatbots are cleared up and their knowledge is expanded.

During a work phase, the teachers develop concepts for the use of AI chatbots in schools, either in groups or individually. They test their prompts on an AI chatbot of their choice or on SchulKI[9], a data protection-compliant AI system specifically tailored for use in schools, with which we have cooperated for some trainings. In a final round in the plenum, the results of this phase are also shared.

Grimm's new Tales (Language Models Unplugged)

In the third round of interventions, we further tested and developed the activity "Grimm's New Fairy Tales" for language models. In this activity, students can understand four concepts of large language models: training a language model by inputting word sequences (K1), the role of randomness in generating new sentences (K2), training and application as different phases in the life cycle of a language model (K3), and the influence of the text corpus on the generated sentences (K4).

We tested the activity on a total of seven interventions during the 2nd and 3rd round. In addition, there were some tests by teachers at schools with school students. We presented initial results and a didactic analysis of the activity at INFOS 2023, the biennial conference of the Didactics of Computer Science group of the German Informatics Society (GI). On our interactively designed poster, interested colleagues were able to insert a sentence into the language model and generate new sentences. For this, we received the Audience Award for the best poster.[10]

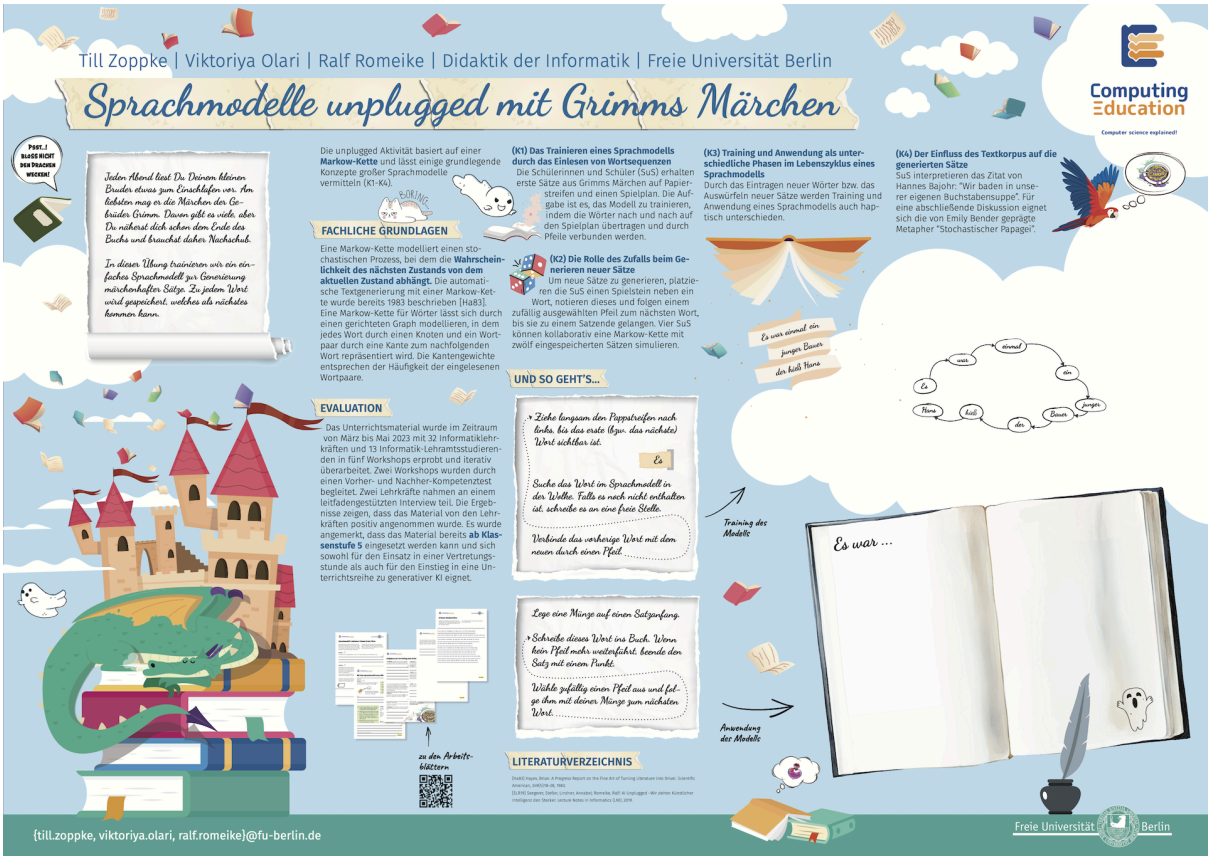


Figure 11: poster submission for the INFOS 2023, documenting the activity "Grimm's new Tales"

[9] https://schulki.de/
[10] https://train-dl.eu/en/news/detail/traindl-team-wins-best-poster-award-at-the-infos-2023

The revised version of the materials, which we completed at the end of the project and presented at a teacher training in Potsdam in March 2024, has several changes compared to the 2023 version:

- The game instructions have been made clearer and more explicit, and partly simplified.
- A total of 12 sentences were selected that have common words and form a word graph with intersections or loops.
- The layout has been revised.

In addition, during one of the weekend workshops, the activity was tested as an introduction, in which the learners are presented with a board with hidden moderation cards. On the back of these cards is a sentence from Grimm's fairy tales, one word per card. The participants must now guess the next word by calling out their guesses aloud. If the guess is correct, the card is uncovered. The teacher can then ask situational questions to the participants, for example why some words are easy to guess and others are difficult, and address grammar, semantics, and the number of possible continuations. This exercise can also make a reference to auto-completion as an informatics phenomenon familiar to students from their smartphones.

To assess that the students have learned the intended technical concepts of the activity such as word prediction, training and stochastic generation, we have developed a gap text. It has to be taken into account though that the test by itself is not sufficient to wrap up the lesson on language models.

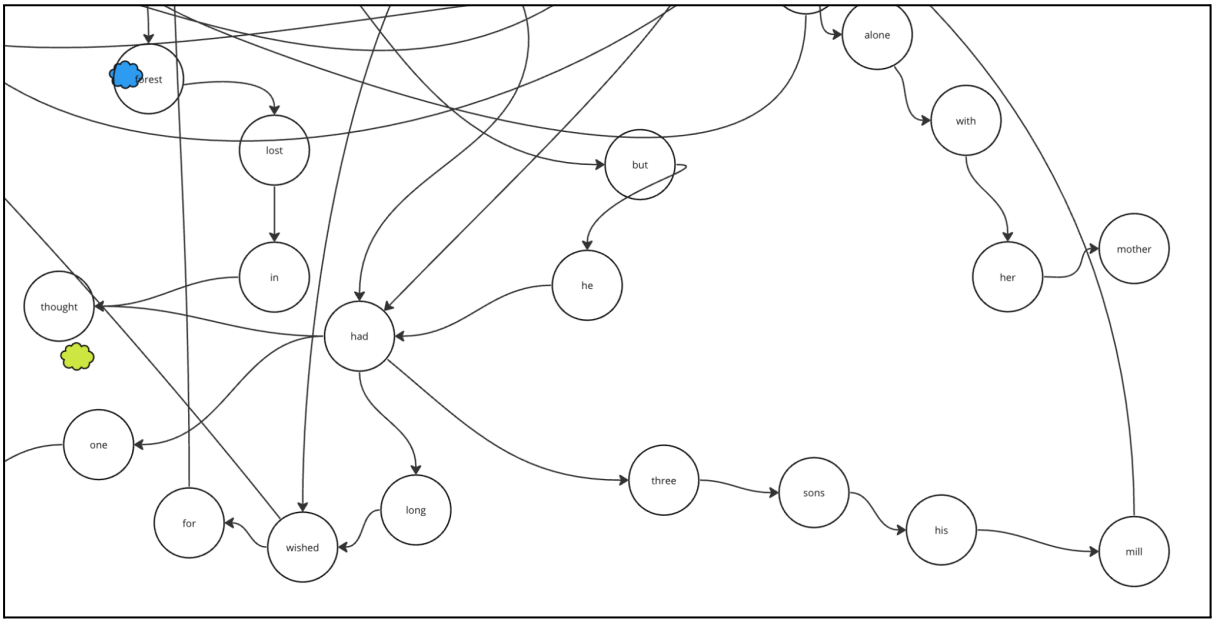


Figure 12. Section from a language graph collaboratively created in Miro during a plugged session

Also, a plugged version of the activity was created based on Miro Board. In this version, the phase where each participant inserts two sentences into their own plan is eliminated. Instead, there is a common plan in which participants insert sentences into the graph structure. In order to keep the structure clear, we have limited the model to eight

sentences. In a second phase, each participant then generates a sentence from the model. Finally the results are presented in the plenary and the process is reflected.

Training Formats

The training formats of the interventions differ depending on the target group and the time frame. In the third round of interventions, we placed special emphasis on the two weekend workshops, one for CS and one for all teachers. We also used the shorter formats (between 2.5 hours and 7 hours), as we did in the first two rounds of interventions, as this is the typical time frame for teacher training, and we were able to use time slots in established events to try out individual activities. The seminar at Freie Universität Berlin stands out among the interventions as it took place over a total of 13 dates during one semester. We present the training formats in detail below in the sections on the individual training sessions.

Evaluation of the effectiveness of the training sessions

The effectiveness of the training sessions was evaluated by the University of Potsdam (UP) as in the previous intervention rounds. For this purpose, the UP questionnaire was reviewed together in a joint preparation meeting and tailored to the intervention. For example, the demographic questions were adapted to the target group of the workshop (type of school, subjects). We also integrated the content questions on the respective activities that were previously in a separate survey into one survey from the UP. The data collected was analyzed by the UP, as reported in the Deliverable D4.6. Furthermore, we conducted individual interviews. The evaluations of the interviews can be found in the Deliverable D4.6.

Two seminars were additionally informally assessed by WP2, in order to pilot our teaching materials and further develop the concept of the teacher training:

- For the seminar "Selected Topics in Didactics of Computer Science," a self-assessment and knowledge test was conducted. The results are included in this report.
- The intervention on October 18th, 2023 in Berlin was evaluated through a questionnaire provided by the organizer. Results and evaluation can also be found in this report.

Implementation of the Training Sessions

Table 1 presents the workshops conducted during the third round of interventions focusing on all teachers (CS teachers, STEAM, humanities, languages and sport teachers). Primary school teachers were also included in the third round of interventions. The interventions with them are reported by WP3 in the deliverable D3.4.. In total, 74 CS teachers and 95 in-service STEAM, languages and sport teachers participated in the third round of interventions.

Table 1. Overview of the training sessions during the second intervention round.

Date	City	Country	Format	Participants
25.04. - 18.07.2023	Berlin	Germany	University Seminar (13 Sessions, 2:15 hrs each)	15 CS pre-service teachers
23.08.2023	Düsseldorf	Germany	4:00 hrs	12 Secondary school teacher-trainers
05.09.2023	Berlin	Germany	1:30 hrs	31 STEAM teachers
18.09.2023	Berlin	Germany	4:00 hrs	27 CS teachers
18.10.2023	Berlin	Germany	1:15 hrs	5 STEAM teachers
23.10.2023	Graz	Austria	3:00 hrs	12 secondary school teachers
24.11. - 26.11.2023	Zeitz	Germany	2.5 days	24 CS teachers
09.01.2024	Vilnius	Lithuania	3:00 hrs	15 secondary school teachers
26.01. - 28.01.2024	Meißen	Germany	2.5 days	23 secondary school teachers

In the following sections, we present the details of the implementation of the training sessions..

25.04. - 18.07.2023 | Berlin | CS Pre-Service Teachers

The seminar "Selected Topics in Didactics of Computer Science" takes place every summer term and is aimed at students in the Master's program for computer science education. In the summer semester of 2023, the seminar was conducted with topics from the field of artificial intelligence. Additionally, the seminar schedule was supplemented with topics on data literacy and AI ethics. Following the recommendations of the stakeholders to focus on in-service teachers and the joint decision by the TrainDL project partners to prioritize the resources of WP4, the seminar was dedicated for the development and evaluation of the teaching and learning activities on AI and data literacy. As in the first round of interventions, we assessed them informally by carrying out a self-assessment and a knowledge test.

Format of a typical seminar session

Following the Dagstuhl Triangle, which we used as a guide in all of our training sessions, the students were first asked to design teaching materials on the technological, socio-cultural,

and applied perspectives of AI and data literacy. This included familiarizing themselves with the topic, researching, adapting and possibly creating their own teaching materials. In the second step, following the didactic doubledecker approach, the students evaluated the teaching materials in a 90-minute lesson. For the lesson, the lead students took on the role of teachers, while all other students and the teachers took on the role of students. In order to keep the lesson challenging, we set the target learner group for the activities as an advanced 12th grade course. After the demonstration lesson, the participants who conducted the lesson reflect on their teaching. This includes presenting their didactic analysis of the subject, their lesson concept, and how they were able to implement their planning and learning objectives. The instructors then resume their role as moderators and lead a seminar discussion, where the participants who experienced the lesson as students share their observations. The seminar session concludes with a summary from the instructors. Table 2 demonstrates the typical structure of a seminar session.

Table 2. Schedule of a typical seminar session in the seminar "Selected Topics in Computer Science Didactics", Summer Semester 2023, Freie Universität Berlin.

Time	**Contents**
	The participants who conduct the lesson prepare the room. Other participants arrive.
9:15	Demonstration lesson on an AI or Data Literacy topic, conducted by two participants
10:45	Break
11:00	The participants who had conducted the lesson present their concept and reflect on the lesson.
11:20	The instructors moderate a discussion, in which the other participants share their observations and suggest alternatives.
11:40	Summary and Conclusion
11:45	End of the seminar session

Seminar Schedule

In terms of content knowledge, participants could select a lesson topic from a total of 10 topics from 4 subject areas: Classical AI, Machine Learning, Data Literacy and AI Ethics. Table 3 outlines the seminar schedule.

Table 3. Seminar schedule of "Selected Topics in Computer Science Didactics, Summer Semester 2023, Freie Universität Berlin.

Date	**Contents**
25.04.	Introduction and welcome to the seminar, Organizational matters, assignment of topics

02.05.	Lesson planning, selection of activities
09.05.	Introduction to AI with an online course Working session: Planning for an introductory lesson on AI
16.05.	Prototype lesson: Language models with Grimm's new tales
23.05.	Lesson 1: Apply and recognize machine learning systems
30.05.	Lesson 2: Intelligent game-playing agents
06.06.	Lesson 3: Knowledge-based AI
13.06.	Lesson 4: Understanding Machine Learning
20.06.	Lesson 5: ChatGPT and how it was trained
27.06.	Lesson 6: Looking into the crystal ball - Prediction models Interim evaluation
04.07.	Lesson 7: Discover hidden secrets - Data Life Cycle and exploratory data analysis
11.07.	Lesson 8: Ethik: Responsible AI in Schools
18.07.	Conclusion and Feedback

Implementation Results and Feedback

In this section, we present the results of the intervention. We begin with an analysis of the participants' work results and learning products. This is followed by a subsection on the evaluation of the self-assessment and knowledge test. Then we discuss the results of the teaching evaluation.

Participant's Work Results

All 15 participants designed a lesson and conducted it in the seminar. In the majority of lesson concepts, students used self-created learning materials. This was also due to the fact that there is often no suitable teaching material published for many topics.

The quality of the lessons varied. In addition to the concept, the success of the lesson also depended on the didactic competence and teaching experience (some of the participants had already completed their internship semester or taught as substitute teachers alongside their studies). Out of 15 participants, 14 participants successfully completed the seminar and submitted written documentations.

In the final session of the seminar, the participants were asked to design a 7h lesson series on AI and data literacy for grade 11 based on the experiences they gained in the entire seminar. The results are presented in table 4. It is noticeable that data literacy is considered to varying degrees. One group has decided to completely remove data literacy from the teaching unit, stating that the topic of artificial intelligence is so comprehensive and important that all seven available lessons should be dedicated to it, and that data

literacy could be addressed in another context, for example in connection with databases. The other groups used one or two double lessons for data literacy.

Table 4: Tabular representation of the series of lessons on AI and data literacy designed by students in the final session of the seminar.

Group 1	Group 2	Group 3
1. Data Literacy 2. AI Definitions and Examples 3. Intelligent Game-playing agents 4. Machine Learning (I) 5. Machine Learning (II) 6. Recap (Examples, test) 7. AI Ethics (Bias, privacy)	1. Introduction (AI in everyday life) 2. Data Lifecycle I: survey design 3. Data Lifecycle II: Cleaning, visualization 4. Machine Learning 5. Clustering 6. Prediction 7. AI Bias	1. AI in everyday life 2. Classical AI vs. Machine Learning 3. Supervised learning 4. Reinforcement learning 5. Neural Networks 6. Language Models 7. AI Ethics

Several teaching materials that were designed and tested during the seminar were further developed by the WP2 and used in the subsequent teacher training sessions (e.g., Activity "Food Quality Control").

Self-assessment and knowledge test

A few items were added to the self-assessment and knowledge test compared to the first round of interventions in order to cover additional topics from the field of AI and data literacy. Out of the 15 participants in the seminar, 12 completed both the pre and post survey in full. We evaluated the self-assessment and the knowledge test both per person and for individual items.

Figure 13. The results of self-confidence before and after participation in the seminar for each participant

The self-assessment shows that participants' self-confidence in their own expertise has increased, as shown in Figure 13. Only individual #3 shows lower values in the post-test than in the pre-test, and also the lowest value of the entire group.

The mean value for self confidence increased from 2.90 (slightly below the value of the neutral option) to 3.69, as shown in Figure 14. The largest delta was recorded by person 8 with an increase from 2.12 in the pre-survey (the lowest value in the field) to 3.92.

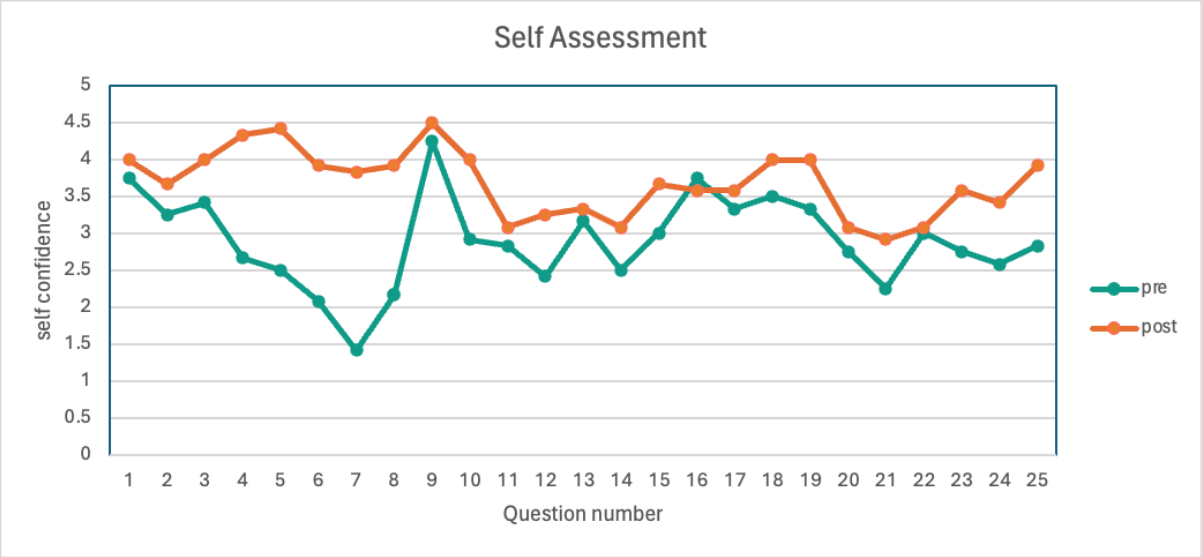


Figure 14. The results of self-confidence before and after participation in the seminar for each question.

Broken down by individual items, an increase in self-confidence can be measured in all items except #16. This item describes the confidence in being able to explain the phase of modeling in the data lifecycle. The decrease from 3.75 to 3.58 is, however, small. It can be observed for all the items that correspond to the self-confidence to explain individual phases of the data lifecycle (items #13 - #22) the increase is low, with an average of 3.03 to 3.43.

The strongest average increase can be seen for items 4-8. These are formulated as "I can...

... explain the principle of supervised learning." (+1.67)

"... explain the principle of unsupervised learning." (+1.92)

"... select a suitable AI method for typical applications, such as classification." (+1.83)

"... interpret a confusion matrix." (+2.42)

"... explain the difference between numerical and categorical data." (+1.75)

Item 9 (“I can explain the difference between correlation and causation.”) shows the highest absolute self-confidence in both the pre- and post-survey.

Figure 15. The results of the knowledge test before and after participation in the seminar for each participant.

In the knowledge test, it can be observed that the students consistently achieve better results in the post-test than in the pre-test (average increase of +0.21), thus indicating an increase in knowledge. Person #6, however, performs worse in the post-test (-0.05). For Person #3, the performance remained almost constant (-0.005). This corresponds to their lower self-assessment at the end of the seminar. However, looking at the raw data, it is noticeable that in the post-test, the person ticked the option "I don't know" for eight tasks, compared to five times in the pre-test. Person #9 achieved the lowest score by far in the pre-test. A look at the raw data shows that this is due to the fact that they chose the option "I don't know" a total of nine times. This could be related to the fact that before the pre-test, we verbally informed the participants that there was an option "I don't know" for every question, and person #9 felt encouraged to choose this option if there was uncertainty about what the correct answer was. The median among all individuals of the time spent for the surveys is 16:17 minutes in the pre-test and 12:26 in the post-test.

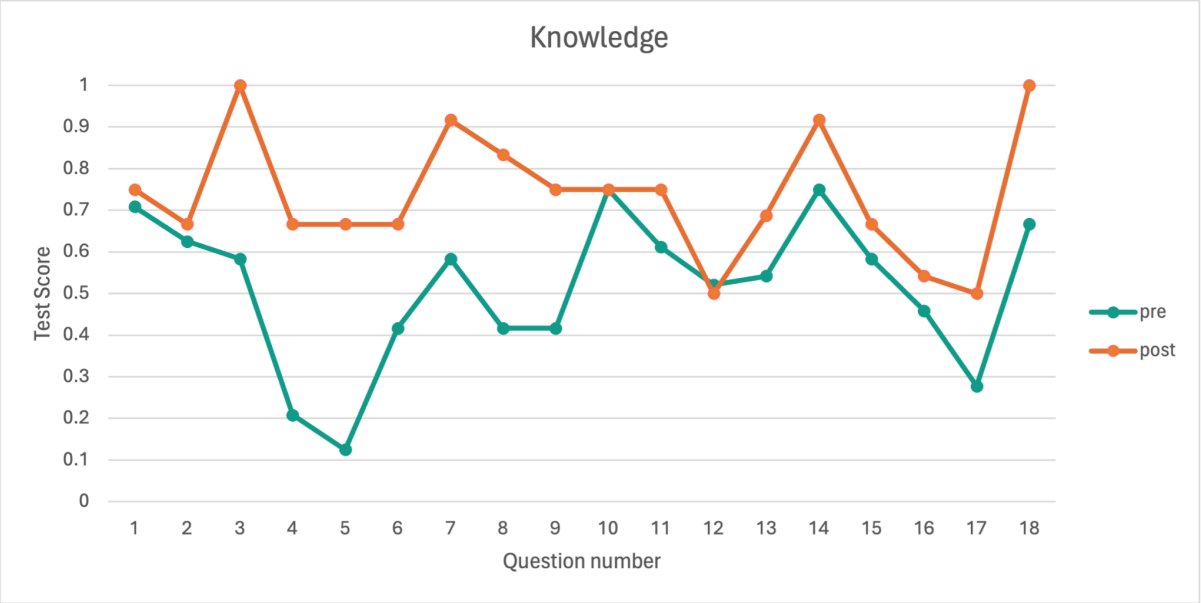


Figure 16. The results of the knowledge test before and after participation in the seminar for each question.

Analyzing the data from the perspective of the individual items, there is a significant increase of more than +0.4 points for four items:

- Item 3: What is the relationship between machine learning and AI? (Correct answer: Machine learning is a subfield of AI.)
- Item 4: Classical AI systems are particularly suitable when... (Correct answers: ... the problem space can be clearly described, ... knowledge from experts is available.)
- Item 5: In unsupervised learning... (Correct answers: ... unlabeled data is used as input. The method identifies similarities and patterns in the input data, for example to group the data or find outliers, ... the programmer sometimes has an influence on the results of the learning process, for example by determining the number of clusters.)
- Item 6: Which method is suitable for predicting the price of a house based on the number of rooms and location? (Correct answer: Regression)

These are all topics that were covered in the introductory sessions at the beginning of the seminar and then were deepened in lessons designed by the students. Item 18 (increase from +0.33 to the maximum value of 1.0), which addresses the topic of AI ethics, should also be emphasized.

There were minor changes with a delta of less than +0.05 for the four items:

- Item 1: What are indicators for an AI system? (The system follows a strategy to solve a task. / The system improves based on experience.)
- Item 2: Which statement applies to weak AI systems? Weak AI systems can... (solve certain well-defined tasks. / learn from training data to solve specific problems.)
- Item 10: When describing a dataset containing data about students from your hometown, you notice: "The higher the education level of the parents, the better the grades of the child." What type of relationship is this statement referring to? (Correlation)

- Item 12: You want to train a model to predict grades based on student data. You don't have time to create your own dataset, so you search the internet for a suitable dataset. What should you consider when choosing the dataset? (4 correct options out of a total of 5 possibilities.)

Regarding items 1 and 2, it can be noted that these topics from introductory sessions on AI were not further explored in the seminar. On the other hand, items 10 and 12 were not addressed in introductory sessions, but were covered in the two demonstration sessions on the subject of data science.

Finally, we also check the data for correlations. The results are displayed in table 5. A moderate correlation of 0.47 was found between the knowledge measured per item in the pre- and the post-test, signaling that an item where participants scored high in the pre-tests is likely to generate high scores in the post-test as well. Another moderate correlation of 0.40 exists between a person's self assessment and their knowledge, as measured in the post-survey. This means that the participants' ability to self-assess their capabilities slightly increased from a correlation of 0.34 in the pre-survey.

Table 5: Tabular representation of the series of lessons on AI and data literacy designed by students in the final session of the seminar.

Assumption	**Axis A**	**Axis B**	**Correlation coefficient**[11]
People with a high confidence before have also a high confidence after the seminar	self assessment per person (pre)	self assessment per person (post)	0.19
People with high knowledge before have also a high knowledge after the seminar	Knowledge per person (pre)	Knowledge per person (post)	0.36
Before the seminar, people with a high confidence have also a high knowledge	self assessment per person (pre)	Knowledge per person (pre)	0.34
After the seminar, people with a high confidence have also a high knowledge	self assessment per person (post)	Knowledge per person (post)	0.40
People who gained confidence in the seminar also gained knowledge	delta in self assessment per person	delta in knowledge per person	0.22
In the pre-test, people who spent more time to fill out the survey also scored higher in the knowledge test	Time spent per person (pre)	Knowledge per person (pre)	0.30

[11] Numbers rounded to two decimal places.

In the pre-test, people who spent more time to fill out the survey also scored higher in the knowledge test	Time spent per person (post)	Knowledge per person (post)	0.27
Self assessment items with a high score in the pre-survey also received a high score in the post-surve	Self assessment per item (pre)	Self assessment per item (post)	0.26
Knowledge items with a high score in the pre-test also received a high score in the post-test	knowledge per item (pre)	knowledge per item (post)	0.47

Teaching evaluation by the participants

The teaching evaluation was conducted as a midterm evaluation through an online questionnaire. 5 students filled out the questionnaire. The timing was chosen by the department so that the instructors have space in the seminar to discuss the feedback with the students and incorporate it into their teaching. In the following, we select each five confirming and five critical items from the total of 36 items.

The pace of the seminar was rated as "exactly right" by all five participants who took part in the survey. Regarding the level of difficulty, three participants believed it was “exactly right”, as shown in figure 17 One participant each indicated it was “rather too low” and respectively “rather too high”. The extent of material covered in the lecture is considered “exactly right” by most participants. Only one person stated it was “rather too low”. In these aspects of the seminar concept, it seems that we have adjusted well to the needs of the students.

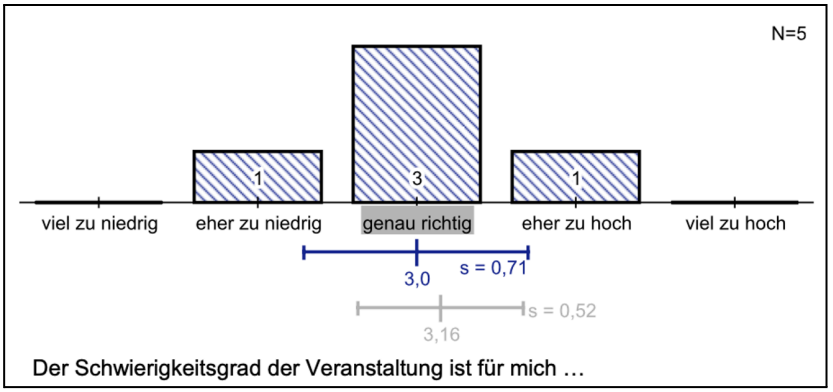


Figure 17. Assessment of the difficulty of the seminar by the students

There is agreement with the statement that the seminar instructors clarify the practical relevance of the course content: two participants indicate that this is “rather true” and three participants that it is “true” (values 5 and 6 on a Likert scale of 1-7). The seminar instructors are also acknowledged to take the students' contributions seriously. This statement was rated as 5 by one person and as 6 and 7 by two people each.

Critically viewed is the time management of the event (see figure 18). The average of the votes cast is "somewhat" exactly in the middle of the scale (4.0). It is also apparently missed that the seminar sessions regularly summarize the most important content of the seminar. Here, the average of the votes cast is 3.4, and thus between "does not apply" and "somewhat applies".

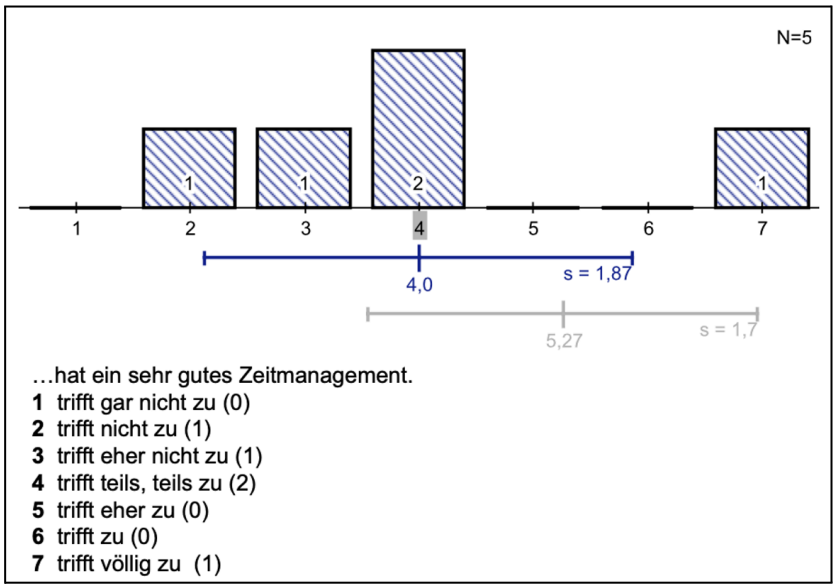


Figure 18. Assessment of the time management during the seminar by the students

Conclusion

In this section we summarize and reflect on the findings above.

Reflection on the seminar concept

In general, the seminar concept proved to be effective. At the beginning of the seminar, it became apparent that the students participating in the seminar had different levels of prior knowledge. This could be due to the fact that approximately half of the participants are studying to become teachers through a cross-entry Master's program. This means that they have already graduated in a subject (such as computer science) and are now studying with a focus on education. Some of the participants had already completed their teaching internship and the corresponding preparation seminar, while others had just finished their Bachelor's degree. These varying levels of prior knowledge led to a great challenge for some to independently research and familiarize themselves with an unknown topic.

Lessons learned

- In addition to the subject-specific and didactic learning objectives, as covered by our in-service teacher trainings, student teachers have professional qualification objectives such as didactic analysis, preparation and implementation of lessons, and subsequent reflection.
- For events that take place over several weeks, it is recommended to repeatedly conduct a survey in the form of a progress test.

- Depending on the level of competency development of the students, prescribed sources for familiarizing oneself with the topic and obligatory consultations before demonstration lessons are recommended.
- The collaborative development and implementation of demonstration lessons (team teaching) was not questioned by the students and appears to be a suitable form of peer learning for student-teachers.
- AI & DL concepts that are covered in the introductory sessions of the seminar and further explored in demonstration lessons are particularly well internalized.
- The students find the reflection on their demonstration lessons and the feedback from the seminar to be helpful.

Reflection on the survey design

We adopted items from the first round of interventions. However, the learning objectives and contents of this seminar were different from other interventions due the more extensive time frame. Nevertheless, the items proved adequate because the general goal of all interventions was the same: to enable teachers to teach Artificial Intelligence and Data Literacy.

The individual topics of the demonstration lessons carried out by the participants were a mere means to an end, so that they could practice working independently on an unknown topic from the field of AI and Data Literacy, to design and conduct a lesson. In the future, an evaluation of these learning objectives could be done through a qualitative analysis of the submitted written assignments, including reflections.

It would have been helpful to conduct a mid-term survey with the identical questions in the fourth session of the seminar in order to assess what the students have learned in the introductory sessions of the seminar. With that, we would have had additional data to support the hypothesis that a dual engagement with topics, in the introductory sessions and in a designed sessions, has a greater learning effect.

23.08.2023 | Düsseldorf | Secondary school teacher trainers

In the workshop conducted in Düsseldorf school teachers participated who were at the same time teacher trainers. Half of the participants were trainers for mathematics at lower secondary level, and 20% were also trainers for music at primary level. There were also two computer science teachers among the participants.

Workshop Schedule

With the organizers, it was agreed as a learning objective that the participants should understand how AI works. AI chatbots (in the following - chat AIs) should be also treated. When planning the course, it was decided to start the morning with personal experiences with chat AIs at school, and to offer stations on AI paradigms in the afternoon. The workshop was evaluated by WP4 with the questionnaire. In addition, a follow-up interview was conducted with one participant. Table 6 outlines the workshop schedule.

Table 6: Schedule of the Düsseldorf Intervention for secondary school teacher trainers

09:00	Internal meeting of the trainer team and information about the new grouping
10:00	Welcome the participants Policy Experimentation Questionnaire (I)
10:15	AI Chatbots at school (act_20, act_21) ● Statistics unplugged: how do you use AI Chatbots at school? ● Think, pair share: experiences with ChatGPT ● Input: how does ChatGPT work? ● Group work: use cases for ChatGPT at school, prompting experiments ● Share the results
12:00	Lunch break
13:00	Paradigms of machine learning (learning stations) ● Beat the robot: classical AI and reinforcement learning (act_03) ● Prototype customer: unsupervised learning (act_04) ● Grimm's new Fairy tales: language models (act_16) ● AI and Music (act_23) ● Material recognition: Image classification with Teachable Machine (act_24)
14:40	Summary: The participants share their experiences from the learning stations.
14:50	Evaluation and Feedback Policy Experimentation Questionnaire (II)
15:00	End

Use cases for AI chatbots at school

In this section we evaluate the use cases for AI chatbots mentioned by the teachers. The majority of participants had not used ChatGPT or had only used it occasionally. We worked with prompts in two stages. First, we formed groups of four people, with two more experienced and two less experienced participants. They wrote example-prompts from their

own practice on moderation cards. These examples were then presented in plenary on a bulletin board. The differentiated composition of the groups was intended to ensure that each group produced results. Teachers without any prior exposure to a chatbot were not able to contribute here, but could learn from their colleagues. Figure 20 demonstrates a worksheet for this task.

A total of 12 prompts were collected and sorted into three categories. The prompts were documented photographically and made available to the participants on an online board.[12] Most of the prompts could be assigned to the category of "lesson preparation". There were examples for conceptual planning ("Plan a class 2 lesson on the topic of Star Singers"), but also for the creation of teaching materials ("Write a fantasy story for an eight-year-old girl"), use cases for research ("Name the 10 most important stages of Kafka's life") were also mentioned. During the discussion, one participant came up with the idea of using a chat AI to differentiate teaching material.

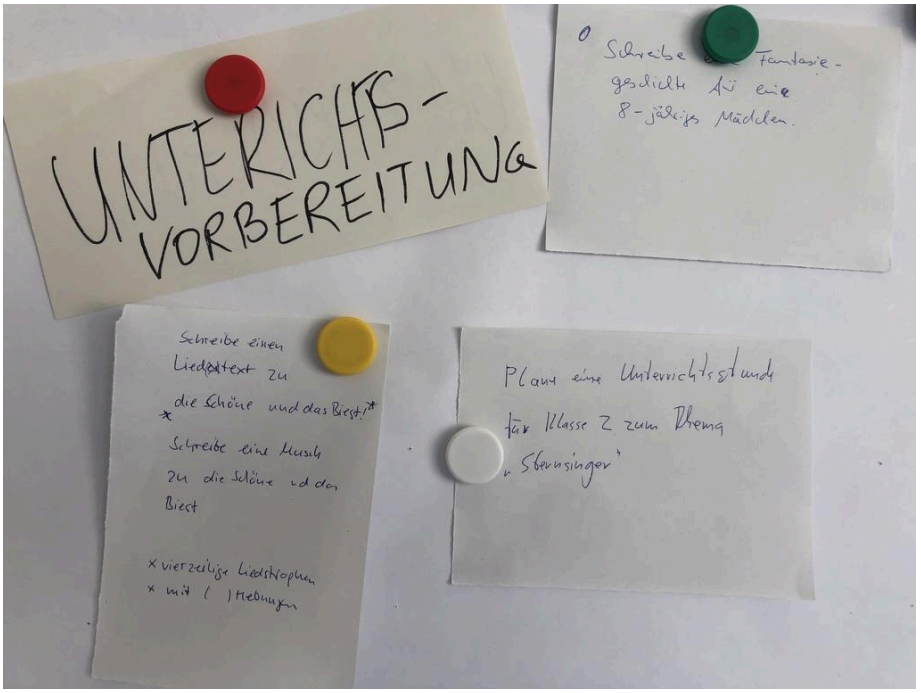


Figure 19: Photo documentation of teachers' prompts for lesson preparation

In a second stage, the participants first experimented with prompts, and then paired up to work on a use case for their teaching. To provide inspiration, we had prepared a collection of prompts, differentiated according to use cases and subjects.

Overall, the participants were very motivated to try out different chatbots like ChatGPT and Fobizz (and continued to do so in the afternoon). One participant expressed the view that teachers should educate students how to use AI chatbots ethically and responsibly.

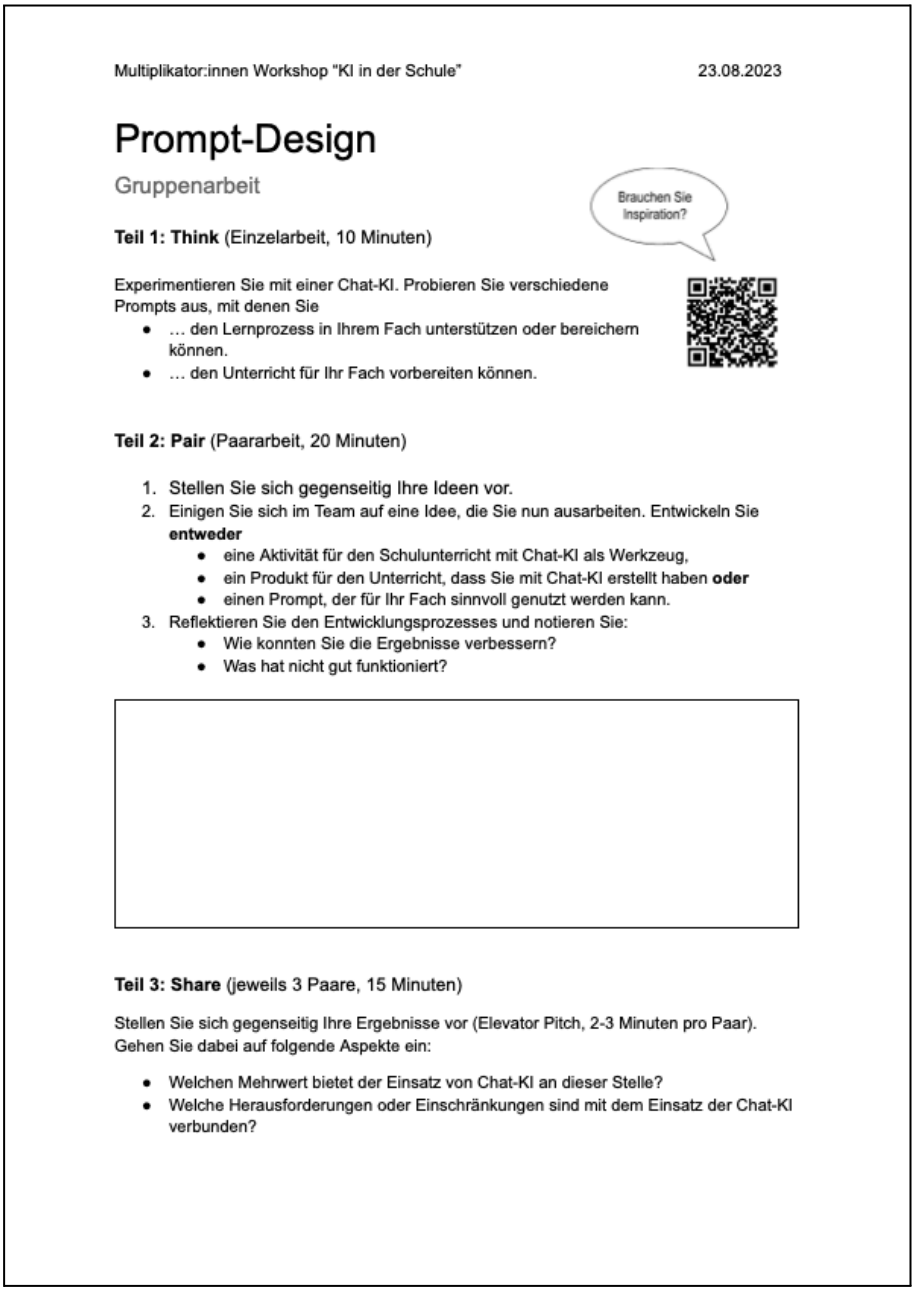
Learning stations

The learning stations in the afternoon were not well received by the participants. Most of them only tried out one or two stations and then continued working with an AI chatbot. We had chosen the stations following the technological perspective of the Dagstuhl Triangle

[12] https://padlet.com/rrettera/sprach-kis-im-schulunterricht-xpian2u68gcunjz7

"How does AI work?" and used activities that had already been successfully tested in previous interventions, such as "Beat the Robot" or "Grimm's New Fairy Tales."

One possible explanation for the lack of interest is that the participants saw no benefit in understanding the technological perspective, because the application perspective was the main focus at the beginning of the workshop.



Multiplikator:innen Workshop "KI in der Schule" 23.08.2023

Prompt-Design

Gruppenarbeit

Brauchen Sie Inspiration?

Teil 1: Think (Einzelarbeit, 10 Minuten)

Experimentieren Sie mit einer Chat-KI. Probieren Sie verschiedene Prompts aus, mit denen Sie

- … den Lernprozess in Ihrem Fach unterstützen oder bereichern können.
- … den Unterricht für Ihr Fach vorbereiten können.

Teil 2: Pair (Paararbeit, 20 Minuten)

1. Stellen Sie sich gegenseitig Ihre Ideen vor.
2. Einigen Sie sich im Team auf eine Idee, die Sie nun ausarbeiten. Entwickeln Sie **entweder**
   - eine Aktivität für den Schulunterricht mit Chat-KI als Werkzeug,
   - ein Produkt für den Unterricht, dass Sie mit Chat-KI erstellt haben **oder**
   - einen Prompt, der für Ihr Fach sinnvoll genutzt werden kann.
3. Reflektieren Sie den Entwicklungsprozesses und notieren Sie:
   - Wie konnten Sie die Ergebnisse verbessern?
   - Was hat nicht gut funktioniert?

Teil 3: Share (jeweils 3 Paare, 15 Minuten)

Stellen Sie sich gegenseitig Ihre Ergebnisse vor (Elevator Pitch, 2-3 Minuten pro Paar). Gehen Sie dabei auf folgende Aspekte ein:

- Welchen Mehrwert bietet der Einsatz von Chat-KI an dieser Stelle?
- Welche Herausforderungen oder Einschränkungen sind mit dem Einsatz der Chat-KI verbunden?

Figure 20: Worksheet on prompt design

Further, it might have been better to set up the stations together with the participants, in order to indicate the new learning setup after lunch.

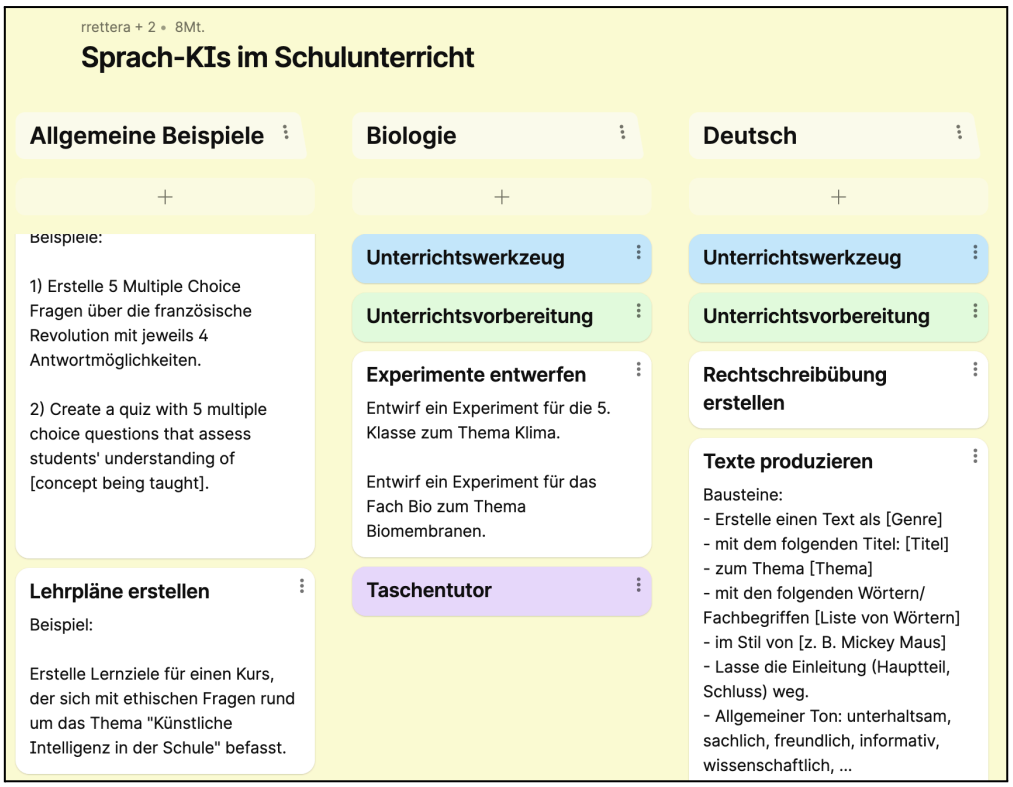


Figure 21: Clipping from the web board with sample prompts for the participants

Finally, we summarize individual feedback on the learning stations. The station on AI & music was well received. The creation of similar materials for other school subjects is conceivable. The referenced online activities are designed as black boxes, allowing for a playful approach to AI rather than an analytical-technical approach. The activity “Beat the robot” was not perceived as an appropriate self-learning activity. In order to use it at a learning station, a corresponding worksheet must be created or (as implemented in this intervention) a supervisor must guide the group work.

In the activity "material control", the participants had difficulties in training an image recognition with sufficient accuracy. This was due to the fact that it proved difficult to take pictures with a neutral background with the webcams of the laptops, as the participants held the material examples in front of the camera with their own hands. The facilitators explained to the participants that the machine learning method does not differentiate between the foreground and background and thus may learn background features instead of the intended items in the foreground.

The activity, in which the participants were able to try out and compare different AI chatbots, was also well received. It introduced alternatives to ChatGPT, such as GDPR-compliant AI tools from Fobizz.[13]

Ideas for future activities

In the context of the workshop, while working with participants and based on their suggestions and their feedback, ideas for enhancements and new activities came up.

[13] https://tools.fobizz.com/

- Since the presentation on the basics of ChatGPT was so well received, we could record it again and produce a video.
- The ministry of education in North Rhine-Westphalia has issued a guideline for the use of text-generating AI systems at schools.[14] This could be used as material for activities addressing AI chatbots. Especially since the participants themselves are multiplicators and in the role of educating teachers, it is important for them to know official documents from their employer regarding AI at schools.
- In preparation for this workshop, we analyzed TikTok videos on AI chatbots and how they can be used to cheat in homeworks. Based on these videos an activity can be designed that puts teachers in the shoes of their students. For example, teachers could be tasked to generate a solution for a typical homework, e.g. a cloze test, a math exercise or an essay by prompting an AI chatbot with the task description. The solutions are then reviewed by a teacher of the respective subject and graded. This may be followed by a reflection on how an AI generated solution can be recognized and which task formats are easily to be solved by an AI chatbot. This leads to the topic "examination culture".
- At the beginning of the collegial exchange, participants can be supported by a template to write down their prompts. This could include the following fields:
    - Description
    - Application area (e.g. teaching, preparation, research)
    - School subject (if applies)
    - Prompt text
- To practice formulating prompts and discover new possibilities, participants could be asked to vary a basic prompt and compare the results. For example, on the topic of "gravity":
    - "What is gravity?" (formulated as a question)
    - "Explain what gravity is." (using an operator)
    - "Explain in two sentences what gravity is." (specify length of length)
    - "Explain gravity to a six-year-old." (specify target group)
    - "Explain gravity in a way that nobody understands." (creative nonsense)
- The station for comparing different AI chatbots can be rethought. A PC can be prepared, with various AI chatbots (ChatGPT 3.5, ChatGPT 4, Google Gemini, Llama2, etc.) opened in their own browser tab, and a user already logged in. On a worksheet, various prompts are given to try out. For each AI chatbot there is a field to take notes or evaluate the response for a given prompt. The list of prompts may include prompts in which the chatbots fail to generate a correct answer, for example "2**4**8" or, to explain hallucination: "You are an investigative journalist. Write an article about Bruno Bertini, the head developer of BERT."

Reflection and feedback

In this section we compile further feedback and observations from the trainers.

- The participants were more interested in AI applications than in AI concepts.

[14] https://www.schulministerium.nrw/textgenerierende-ki

- We initially had difficulties finding volunteers for interviews after the workshop. This can be potentially prevented by personally asking participants during the workshop, if they are willing to volunteer.

05.09.2023 | Berlin | STEAM teachers

The Association for the Promotion of STEM Education (MNU)[15] holds an annual state conference in Berlin. At the 2023 conference, the TrainDL project was invited to conduct a workshop. The venue took place at Freie Universität Berlin.

In the MNU program, workshops are offered for both individual STEAM subjects and cross-curricular subjects. With learning goals from the field of data literacy, we aimed the workshop at teachers from all STEAM subjects. Due to the domain of the Abalone activity, biology teachers may have felt particularly addressed by the workshop. Figure 22

WF_04 **05.09.23 11:30** **90 min**

Viktoriya Olari, Till Zoppke, Ralf Romeike
FU Berlin, Institut für Informatik, AG Didaktik der Informatik

Workshop

KI und Datenanalyse im MINT-Unterricht

Sek 1/2

Der Workshop 'KI und Datenanalyse im MINT-Unterricht' bietet Ihnen eine praxisnahe Einführung in die KI-gestützte Datenanalyse, die in einer Zeit, in der Künstliche Intelligenz immer mehr Anwendungen findet, für Schülerinnen und Schüler eine Schlüsselkompetenz darstellt. Durch den Einsatz von KI-Werkzeugen im MINT-Unterricht können Schülerinnen und Schüler neue, innovative Herangehensweisen an Probleme erlernen.

Anhand von konkreten Fallbeispielen (Case Studies) und ausgewählten Datensätzen lernen Sie, Probleme mit Hilfe von KI-Werkzeugen zu lösen. Die Vorgehensweise orientiert sich dabei am Daten-Lebenszyklus und beinhaltet die Exploration, Analyse und Visualisierung von Daten. Im Schulunterricht ermöglicht der datengetriebene Lösungsansatz den Schülerinnen und Schülern neue Perspektiven auf Fragestellungen der MINT-Fächer.

Bitte installieren Sie das Datenanalyse-Werkzeug Orange3 [1] und bringen Sie Ihren Laptop zum Workshop mit. Programmierkenntnisse sind nicht erforderlich.

Lernziele:

- KI-gestützte Datenanalyse als neues Unterrichtswerkzeug kennenlernen.
- Problemstellungen aus den MINT-Fächern mit dem Tool “Orange3” bearbeiten.

[1] orangedatamining.com/

Figure 22: Invitation to the workshop for STEAM teachers[16]

Workshop schedule

Due to the limited time frame of only 90 minutes, we decided to tailor the workshop entirely to the developed activity on multiple linear regression using the example of Abalones (sea snails). This also included an introduction to the mathematical foundations of linear regression. This method of machine learning is used in the exercise to predict the age of abalones.

An introduction to the topic of AI and data literacy is also included in order to provide context for the workshop. Data science is considered a new fourth pillar of science, making it methodically relevant in the context of a scientific introductory course in upper-level

[15] https://www.mnu.de/
[16] https://www.mnu.de/extern/mitglieder/anmeldung/teilnehmer/DetailsBeitrag.php?Beitrag=4876

education beyond just computer science classes. After a brief welcoming part, the participants were asked to fill out the policy experimentation questionnaire of WP4.

Table 7: Schedule of the STEAM Intervention for secondary school teacher trainers at Freie Universität Berlin

11:30	Welcome the participants Policy Experimentation Questionnaire (I)
11:40	Introduction to AI (act_02) ● Grades 11+ prepare students for university ● Data-centered AI as the fourth pillar of science
11:50	Abalone (Act_25) ● Introduction to the abalone data set ● Excursion: Linear Regression ● Live demonstration: a simple data workflow in Orange
12:15	Abalone: Participants work on their own laptops and solve the worksheets
12:45	Evaluation and Feedback Policy Experimentation Questionnaire (II)
13:00	End

Implementation

The workshop received great interest and was fully booked. We had a total of 31 participants, which exceeded the number of registrations. The introductory presentation was followed with great interest by the participants, who also asked a few interim and follow-up questions. This caused us to fall behind schedule, so that the participants had less time than planned to work on the worksheet and the final evaluation in plenary was canceled.

Figure 23: One of the instructors, answering questions during the training.

The linear regression method was explained on a chalkboard. This setting was deliberately chosen, based on a typical teaching situation from math lessons, in order to familiarize the participants with the mathematical concepts. In technical terms, linear regression is based on the topic of linear functions, which is covered in class 8 in Germany. This connection made it easier for the participants to familiarize themselves with the explanation.

An important learning objective of the workshop was also to familiarize the participants with Orange3 as a tool for analyzing data and its possible application in the classroom which was difficult due to a limited time of the workshop. A slide comparing the features of Orange3 with the features of Excel could help participants to understand the functionality of Orange and how its approach to data differs from the one of Excel.

As an initial example, participants could create a simple data workflow in Orange. For example, they could use linear regression on a data set to predict a person's shoe size based on their height. This would be an intermediate step after the conceptual explanation of linear regression and before creating the prediction model for the age of the abalone.

The task sheet itself was well received by the participants. In particular, the real-world domain of sea-snails was praised as being suitable for students. To understand how a prediction model can be determined using the linear regression method, participants suggested keeping an eye on the relationship between the data and the real world. For example, if the training results in a function with a negative ordinate for the length of the abalone, or generally which units are to be assigned to the individual columns. A slide that illustrates the principle of computer modeling and thus the solution approach chosen for the prediction problem could possibly also help here. Another dedicated slide could summarize the domain knowledge about sea snails: where do they live, what do they feed on, how big and how heavy are they typically. Alternatively, this information could be added to the worksheet. Another piece of technically relevant information suggested by the participants was the criteria that a data set should ideally fulfill in order to create a suitable prediction model with the help of multiple linear regression. This also includes the fact that the independent variables should not correlate with each other. In our example, however, a multicollinearity was given, as width, height and length correlate with each other. This and similar background information could be summarized in an explanatory text at the end of the worksheet.

We summarize the technical and organizational best practices from the training:.

- It was helpful to ask the participants to install Orange before the workshop and bring their laptops when announcing the event.
- At the beginning of the working phase, it should be clarified which participants may not have a laptop with them or do not have Orange3 installed. Pairs can then be formed if necessary, and two people share a laptop.
- USB sticks with the portable version of Orange3 (Mac and Windows) can be prepared so that the installation can be done on site on individual laptops.
- The colleagues from WP4 were already in the room before the start of the workshop and gave the arriving participants handouts with the QR codes for the Policy Experimentation Questionnaire, so that participants already could use the time to fill out the questionnaires. This reduced time that participants spend on filling out questionnaires before the workshop.
- In workshops like this, where there is no time for a round of introductions, it is helpful to ask the participating teachers at the beginning what type of school

(primary or secondary school) and what subjects they teach. This allows the instructors to get an idea of the participants' previous knowledge and interests.

- If there are a sufficient number of volunteers, an interview can also be conducted with several participants. One person can then respond to points mentioned by another person and they can complement each other.

18.09.2023 | Berlin | CS teachers

The Informatiktag (engl: Computer Science Day) Berlin-Brandenburg takes place every year in the first weeks of the school year. It is a joint event organized by the Regional Training Berlin, the specialist group "Computer Science Education in Berlin and Brandenburg" (IBBB) of Gesellschaft für Informatik (GI), and the LISUM. The TrainDL project has already carried out an intervention and a pilot here, in 2022 and 2021 respectively. Figure 24 demonstrates the announcement of the workshop on the event website.

KI-Expedition: mit lernenden Maschinen durch das Land der Daten
Dozenten: Viktoriya Olari, Ralf Romeike, Till Zopke
Leitung: Alexander Dietz

Der Workshop richtet sich an Lehrkräfte, die die Themen Künstliche Intelligenz und Data Literacy in ihren Unterricht integrieren wollen. Texte, Bilder, Zahlen können mit den Methoden der Datenanalyse und der Künstlichen Intelligenz neue Erkenntnisse über die Welt liefern und uns helfen, verschiedene Probleme zu lösen. Nicht umsonst gilt die Fähigkeit, mit Daten planvoll und kritisch zu arbeiten, heute als viertes Standbein der Wissenschaft. Mit generativer KI können zudem kreative Leistungen automatisiert werden. Begleiten Sie uns auf einen Ausflug durch das Land der Daten. Nach einer kurzen Einführung in den Lebenszyklus des Maschinellen Lernens probieren Sie an Stationen Aktivitäten und Unterrichtsmaterialien zu den Themen Bilderkennung, generative KI, Sprachmodelle und intelligente Agenten aus. Sie arbeiten sowohl unplugged mit Papier und Schere als auch mit digitalen Anwendungen wie der Teachable Machine. Beispielsweise schlüpfen Sie in die Rolle eines Data-Scientist, der auf der Basis von medizinischen Bildern ein Neuronales Netz trainiert, das mit Malaria infizierte Blutproben von gesunden unterscheiden kann. Der Informatikunterricht kann dazu beitragen, den Schülerinnen und Schüler wichtige Kompetenzen in Data Literacy und KI zu vermitteln, wie z.B. den Zusammenhang zwischen den Trainingsdaten und der Qualität der Diagnosen. Die entsprechenden Materialien und Handreichungen stellen wir Ihnen zur Verfügung, so dass Sie sie in Ihren Unterricht einbringen können. Der Workshop wird im Rahmen des europäischen Projekts TrainDL angeboten. Dieses Projekt hat zum Ziel, Bildungskonzepte für Datenkompetenzen und Kompetenzen für Künstliche Intelligenz (KI) zu erarbeiten und in der Lehrkräftebildung zu erproben. Auch Lehrkräfte, die bereits an unseren Workshops teilgenommen haben, werden neue Materialien entdecken können.

Figure 24: The announcement of the workshop on the event website[17]

As in previous years, the 2023 event took place at Hochschule für Technik und Wirtschaft (HTW) in Berlin Oberschöneweide. Therefore we were familiar with the organization structure and the audience. The workshop was fully booked. We also allowed additional participants to join. Overall, 27 CS teachers participated in the workshop.

Workshop Schedule

The aim of the intervention was to gather experiences with activities that have already been tested (AI Quiz, Abalone, Malaria/Material Recognition) with CS teaches. In addition, we tested a new activity on Intelligent Agents based on the programming game

[17] https://bildungsserver.berlin-brandenburg.de/informatik/15-informatiktag-berlin-brandenburg-2023

SuperCodingBall.[18] We did not repeat activities that we had tested at the event the year before, in order to offer the participants something new.

The basis for the schedule was our workshop concept for the third round of interventions, based on the method of the didactic double-decker. The workshop was announced as "AI expedition: With learning machines through the land of data", as it comprised independent activities from the fields of AI and Data Literacy. Table 8 summarizes the training schedule.

Table 8: Schedule of the Berlin Intervention fo CS teachers (Informatiktag Berlin Brandenburg, 18.9.2023)

10:45	Welcome the participants Policy Experimentation Questionnaire (I) Introduction round
11:00	AI Quiz (act_20)
11:25	Introduction to AI (act_02)
11:35	Abalone (act_25): Multilinear regression with Orange3
12:45	Lunch break
14:00	SuperCodingBall (act_26): AI agents
14:55	Malaria (act_19) and Material recognition (act_24): Image recognition
15:45	Evaluation and Feedback Policy Experimentation Questionnaire (II)
16:00	End

We had announced the event as a work at learning stations, but after the experiences of the second intervention round, we decided to at least conduct the morning in such a way that all participants work on the same activity. As had proved successful in preceding interventions, the evaluation was carried out with the Policy Experimentation Questionnaire by WP4. The interviews were also conducted directly after the event.

Implementation

In this section we present the impressions of the instructors. In general the participants engaged effectively in the activities.

The AI Quiz was well-received, although the larger-than-anticipated group sizes meant the allotted time for the groups to sort the cards was insufficient. This highlighted the need for clearer task instructions to avoid misunderstandings, such as one group mistakenly sorting the cards based on whether or not the text on the back appeared to be AI-generated to them.

[18] https://www.supercodingball.com/home

The main task of the Abalone activity was creating a data workflow with Orange3. A highlight of the session was when one participant realized that it was odd for edible sea-snails to have body sizes below one millimeter. The instructors used this discovery to pause the working phase. They explained the reason for the low values: to prepare the dataset for machine learning and to reach numbers between 0 and 1, the curators of the dataset had divided all continuous values by 200. This illustrates the importance of domain knowledge for data projects.

The SuperCodingBall activity was only briefly introduced with a life-demo of the programming game. The intention was to test whether the worksheet was suitable for a learning station. However, it turned out that many participants had difficulties implementing even simple game strategies. One reason for this was that the participants' attention was tied to the screen when interacting with the programming environment and they forgot about the worksheet and its guidance. Secondly, many were unfamiliar with the control loop, a central paradigm in programming robots, which consists of an infinite loop in which an agent perceives the world, decides on an action and then executes it.

An ideal solution would therefore be to integrate the contents of the worksheet into the programming game and implement it as an online-only activity. As SuperCodingBall is open source, this would be possible by forking the repository. Alternatively, the online activity could be placed in a sandwich: the participants would start with an exercise on the control loop paradigm (1), then compete in the programming game (2), and finally reflect on the concept of intelligent agents (3).

For the last activity (image classification with Teachable Machine), participants chose between two variants. The worksheet on malaria was significantly more popular than the material control, perhaps because the exercise was already mentioned in the workshop's announcement text, or because it was inconvenient for the participants to take their own training images, as required for the material control.

Feedback by the participants

At the end of the workshop, we asked the participants to get together in groups of three and give feedback on the following questions.

1. What new things have I learned?
2. What topics do I want to know more about?
3. What activities could be suitable for daily school life / for my subject?

We collected the feedback on the whiteboard and documented it photographically (see Figure 25). A tabular overview on the participant feedback is given in Table 9. Below we summarize the points that were mentioned most frequently.

- The tools used in the training were mentioned frequently, both in summary (5x) and individually: Orange (6x), SuperCodingBall (2x) and Teachable Machine (6x).
- However, only three groups stated that a tool was suitable for teaching. Orange (1x) and Teachable Machine (2x) were mentioned.
- 10 mentions can be assigned to information technologies: Data Science (4x), Machine Learning (4x), and Robotics (2x).
- From these mentions, the majority were in response to the question "What topics do I want to know more about?". This can be interpreted as teachers wanting to have a

deeper understanding of the underlying technologies in addition to the teaching material they have already learned.

- 9 responses related to application contexts for Artificial Intelligence or to our activities named after them: Abalone (2x), Soccer (1x), Malaria (1x), and Material Study (1x). Six of these mentions were in response to the question: "What activities could be suitable for school?"

Table 9: Tabular overview of the feedback given by the participants

Label	What new things have I learned?	What topics do I want to know more about?	What activities could be suitable for school?
Tools	„Es gibt niedrig- schwellige Angebote“ „Die Tools“ „Neue Tools kennen gelernt“ „Interessante Tools:“	„Mehr Tools für KI-Projekte im Unterricht (wie die Bsple von heute)“	
Orange3	„Orange3“ „Orange Data mining“ „Orange kennengelernt“ „Orange“	„Orange3“	„Regression & Data Orange“
SupercodingBall	„SuperCodingBall“ (2x)		
Teachable Machine	„Teachable Machine“ (2x)	„Verbindung Teaching Machine & Robotic“ „Hintergrundwissen zur Trainingsphase (Teachable Machine)“	„Teachable Machine“ (2x)
Data Science	„Big Data / Datasets“ „Mit Datasets trainieren und testen: Modellgenauigkeit“ „Daten reinigen“		„Regression & Data Orange“
Machine Learning		„Hintergrundwissen zur Trainingsphase (Teachable Machine)“ „Neuronale Netze“ „Tensorflow“	„Training von KIs“
Robotics		„Verbindung Teaching Machine & Robotic“ „Robotik“	

Application domains	„Abalone = Meeresschnecken“ „Beispiele machen KI greifbar / vorstellbar“	„Überwachung, Militärische Einsätze“	„Pergamon“ „Malaria“ „Materialstudie“ „Abalonen = mal nicht Mathe“ „Kontext für Einstieg in Datenbanken -> Data Literacy (Abalone) “Fußball”
other	„Kriterienorientierung“	“Blockprogrammierung” „KI-gestützte Bildbearbeitung”	

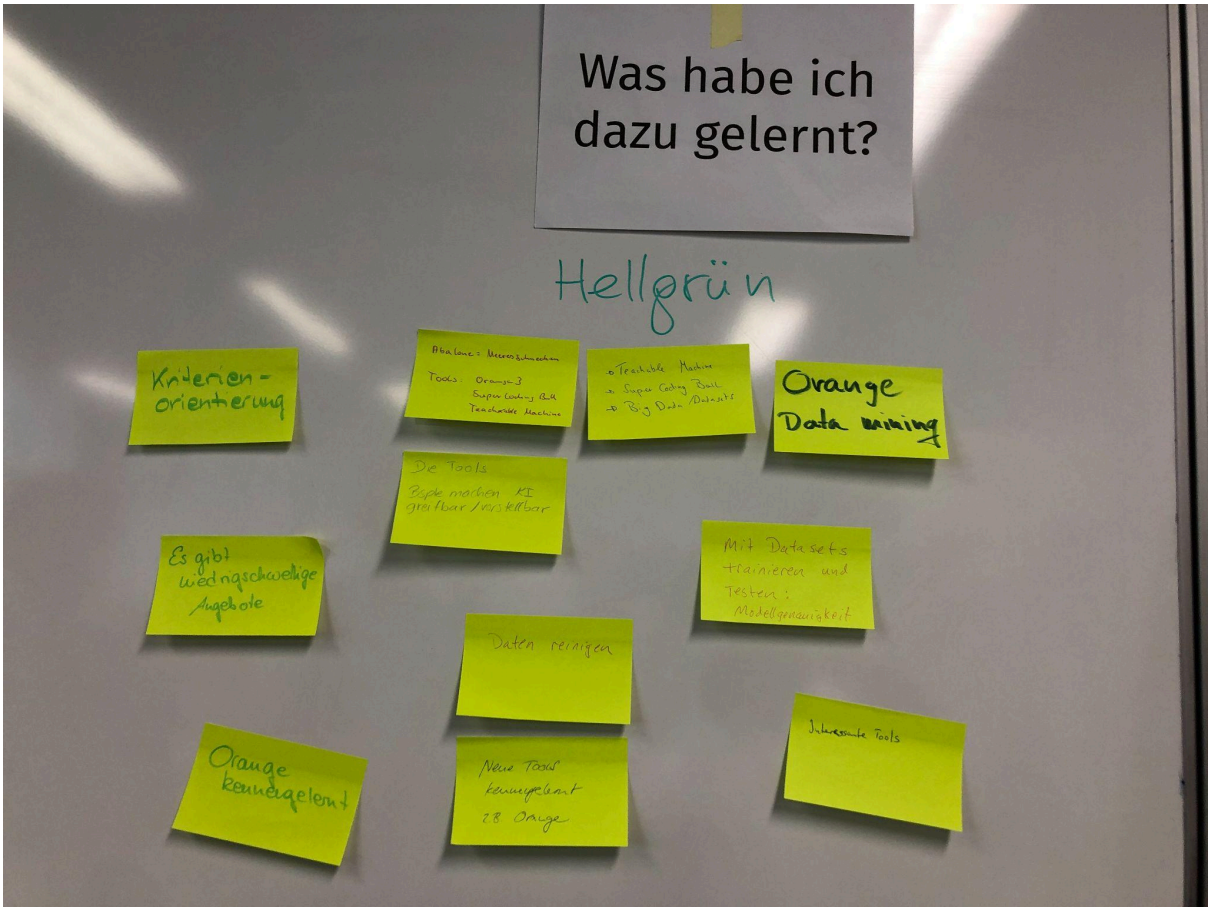


Figure 25: participant feedback at the Berlin Informatiktag, 18.09.2023

18.10.2023 | Berlin | STEAM teachers

The event was organized by the Berlin Senate and targeted teachers on the topic "AI in schools". It was scheduled for a total of 4 hours. After a keynote speech, two blocks of eight workshops each took place in parallel, each workshop of 75 minutes. Figure 26 illustrates the invitation to the event.

Figure 26: Invitation to the training on 18.10.2023

Workshop Schedule

At the request of the Senate, the Freie Universität contributed a workshop titled "Grimm's New Tales - Language Models Unplugged" that was scheduled for the first workshop slot. The target audience should be computer science teachers and technically interested educators. Table 10 presents the schedule of the training session.

We decided to use the workshop to test our unplugged activity on language models. As a start, the instructor played a game with the participants. Cards with words, attached to a board so that the words were covered, had to be guessed one by one. They were then uncovered by the instructor and eventually formed the first sentence of a fairytale. The goal of this activity was to teach the learners how a large language model is typically trained.

The time frame also allowed for an introductory presentation on the topic of generative AI with a focus on large language models. In a discussion, the participants discussed which skills students should have in regards to language models according to the teachers. The workshop ended with an evaluation and reflection.

Table 10: Schedule of the training session that took place at 18.10.2023

14:45 - 14:55	Welcome the participants
14:55 - 15:00	Activity: guessing a sentence
15:00 - 15:10	Presentation: introduction to generative AI (act_15)
15:10 - 15:15	Activity: what should students know about AI chatbots?

15:15 - 15:45	Activity: Grimm's new tales (act_16)
15:45 - 16:00	Reflection, evaluation and feedback

Implementation

The workshop was planned for 12-20 participants. Two days before the workshop 8 teachers had registered. Only 5 participants showed up. Because of a delay in the previous lecture, the workshop had to start 15 minutes later. The participants had the following background from computer science, religious education, ethics.

Overall, the feedback for the workshop was consistently positive. The guessing game with the covered cards on the board worked well. While playing we reflected on how the participants came up with their ideas and could touch on concepts to be explored later.

- Before the start, a participant wanted to know what was on the back of the cards: words, syllables or letters? For a language model it is a design choice. A prototypical large language model (LLM) is based on tokens.
- At some points the available next words are limited because of the syntax (grammar) of the sentence.
- At other points the available next words are limited because of the semantics of the sentence.
- Participants asked whether they also should guess the punctuation => LLMs treat punctuation as tokens. In this game we neglect them for simplification.
- We also discussed how ChatGPT "decides" which word to produce next. => In general, always the most probable one, though this is controlled by the "temperature" parameter.

Overall, the feedback was that the activity can be included in classroom teaching in this way. A didactic guide can describe impulses on how the teacher can address the concepts presented in the following unplugged exercise on Grimm's fairy tales. A connection can be made to how a smartphone predicts words when typing a message.

Also successful was the phase in which participants were asked to brainstorm what students should know about language models, such as:

1. ChatGPT has its "knowledge" from the internet.
2. ChatGPT actually doesn't know anything, it is trained with data from the internet.
3. The information from ChatGPT may not be accurate.
4. Students should verify what ChatGPT outputs.
5. ChatGPT can be used for misinformation and propaganda.
6. ChatGPT reflects the past.

It is noticeable that the teachers use ChatGPT as a synonym for language models / AI chatbots. The competences suggested by the participants represent an impressive number of aspects of AI chatbots. All aspects of the Dagstuhl Triangle are covered: the technical (1,2,3,6), the application-related (4) and the societal perspective (5). The following technical terms can be mapped to the statements: training corpus (1,2), anthropomorphization (2), hallucination (3), information competence (4), manipulation (5), and cutoff date (6).

The main activity, the unplugged exercise with Grimm's fairy tales on language models, was well received by the participants. One participant did not understand how to solve task 1 (reading the sentence already stored in the model). Apparently, the graph structure is more intuitive for CS teachers than for people without a CS background. As some participants needed longer to complete their models, the other people had some idle time. This was offset by participants discussing the exercise and already reading the task for the collaborative generation of sentences.

The simulation of randomness when generating new sentences was done "in the head" or with a counting rhyme. The coin that was used to indicate who is currently taking a turn proved to be valuable (in an earlier version of the activity, this rule was not included). Participants asked, as previously in other interventions, when a sentence is considered complete and a period should be placed. This will be taken into account in the revision of the material, when we will introduce a dedicated “stop” field.

To assess the activity, the participants remained in the room even though the break officially had started. We checked which competencies were promoted by the activity. The teachers argued that (1,2) partially were covered (the Grimm model is trained with fairy tales, but not from the internet). They also mentioned they got technical insight that cause the phenomena covered by competences 4 and 5. We also reviewed the competences that we identified to be covered by the activity, as published in our INFOS poster.

Workshop evaluation by the organizer

The workshops carried out at the event were assessed by the organizer. A questionnaire was sent in an email to the participants asking them to evaluate the workshop. This resulted in a response rate of 80% (4 out of 5 participants completed the questionnaire). The results (see Table 11) confirmed the positive feedback from the workshop.

Table 11: results from the workshop evaluation by the organizer (1 = full agree, 4 = full disagree)

Please indicate how much you agree with the statements.	**Average answer**
The event followed a comprehensible concept.	1
Active participant methods were used.	1
I was able to relate theoretical inputs to my practical work.	1.75
The event offered opportunities to reflect on my professional practice and actions.	1.25
Experiential knowledge and suggestions from the participants could be exchanged.	1
The working atmosphere was conducive to learning.	1
Impulses from the event can be integrated into my professional actions.	1.5
I have benefited in terms of content from the training.	1

23.10.2023 | Graz | Secondary school teachers

The intervention in Graz took place at the University in the context of a teacher training event and was conducted by a TrainDL project member from OCG. The time slot provided by the organizer was 3.5hrs including a break. A total of 12 secondary teachers attended. One of these teachers taught computer science, the others had majors different from CS.

Workshop Schedule

The workshop followed the same schedule as the successful intervention from 17.05.2023 that took place at the same venue and was also conducted by OCG. We provide the timeline here as a sketch. Details on the individual phases can be found in the corresponding section in D2.6 - Report on the second round of interventions.

Table 12: Schedule of the training session that took place at 18.10.2023

Time	**Contents**
5’	Introduction of the TrainDL project
5’	Policy experimentation questionnaire (I)
5’	Getting to know the participants
5‘	Motivation & Activation: AI Quiz (Photos, act_01)
20’	Introduction to AI (presentation, act_02)
3 x 45’ 15’ break	Learning stations: ● Data analysis with Orange3 (act_07, act_18) ● Grimm’s new tales (act_16) ● Teachable Machine with Malaria (act_19) After stations 1&2 there is a 15-minute break.
20’	Affinity Mapping: Teachers share the use of AI and data science in their own subjects.
5’	Present the results of the Affinity Mapping
10’	Feedback & Outlook
10’	Policy experimentation questionnaire (II)

Participant Feedback

The participants felt that the contents of the activities met their level of prior knowledge, and the training was structured. The practical aspect that the participants were able to experience teaching material from the perspective of students (the didactic doubledecker) was emphasized. Positively mentioned were the unplugged activity on language models (Grimm’s new tales) and the activity based on Teachable Machine (Malaria). The teachers

wished the training had lasted longer and suggested a blended learning concept for future training courses, combining practical exercises with online materials.

24.11. - 26.11.2023 | Zeitz | CS teachers

After the weekend workshop for computer science teachers in Heidelberg (November 2022, first round of interventions), our second weekend workshop took place in Zeitz in November 2023. As the organizers, we were free to decide on the organizational framework and content structure.

Figure 27: Conference room at Alte Nudelfabrik, Zeitz

As the workshop was aimed at CS teachers from all over Germany, good accessibility was a criterion for the choice of location. We initially enquired again at the ISG Hotel Heidelberg, which we knew from 2022. As the 2022 training course was perceived as very successful, a repeat at the same location would have given us the opportunity to establish a tradition. However, the hotel was fully booked, so we had to look for other conference venues and decided in favor of the old pasta factory in Zeitz (Figure 27)[19]. Zeitz is easily accessible from Berlin and other places in Germany, and the Nudelfabrik is within walking distance of the railway station.

[19] https://nudelfabrik.eu/

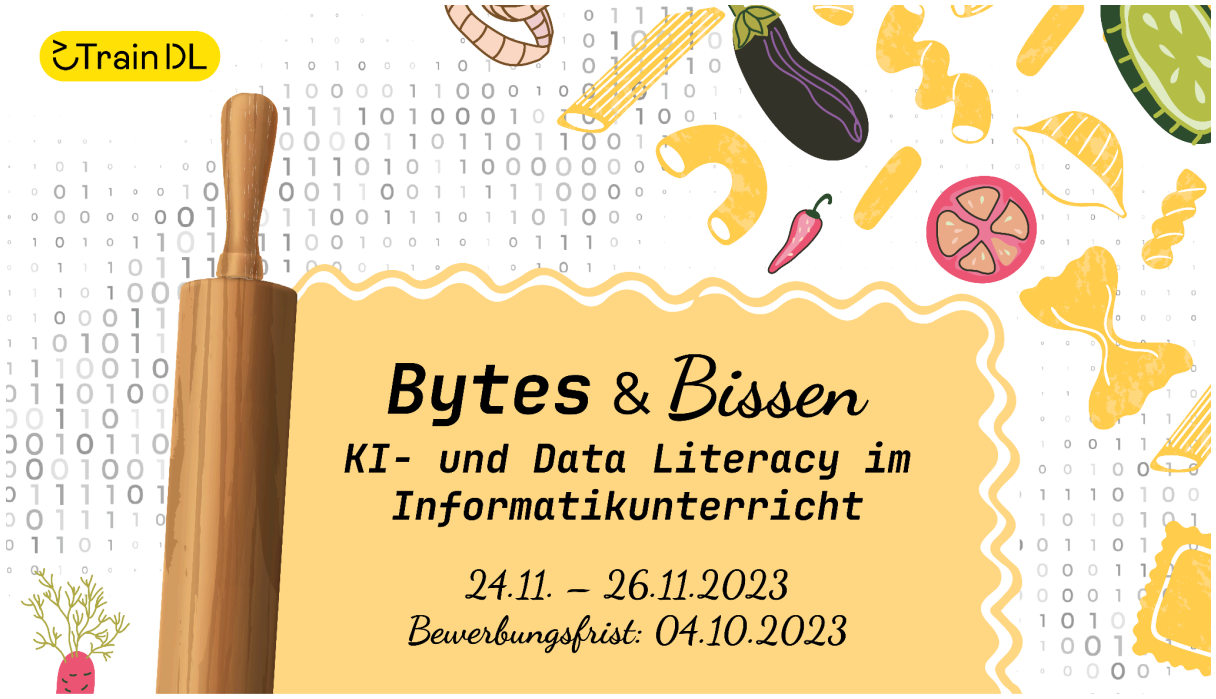


Figure 28: Invitation flyer for the CS teacher training in Zeitz

Workshop Schedule

We reused the tried-and-tested structure from the workshop Heidelberg for the schedule. To cover all phases of the data life cycle, we again planned a data collection and evaluation part. Inspired by the history of the location as a pasta factory, we selected food and eating as real-world domain and thematic framework for the workshop (see Figure 28). The title of the workshop (Bytes & Bissen, engl.: Bytes and bites) was also derived from this. It was also fitting that a pizza oven was set up in the kitchen, which allowed us to schedule pizza baking as a together-action for Saturday evening.

We designed further activities from the food area for the workshop.

- The Generative Cocktail Producer (GCP): an activity in which the participants create a knowledge base and generate cocktail recipes. This activity, scheduled for Friday evening, led into the social part of the evening, where participants could try out their AI-generated recipes and mix fruit cocktails.
- An expert presentation on creative artificial intelligence that explored the theory of food pairing - which ingredients can be combined to recipes perceived as delicious. Beneath the biochemical analysis of flavors the data analysis of millions of recipes is a key research method to gain insights in this field of science. This provided the theoretical background to the "Generative Cocktail Producer" activity. Our expert for the talk was an AI researcher specialized in creative AI.
- Food control - an activity that highlights causes of algorithmic bias in image classification on the example ripe and overripe bananas.

For Sunday afternoon, to reflect on the workshop, we scheduled an activity in which the participants discussed which data skills students need to understand and develop AI applications.

The evaluation of the intervention was carried out using the proven instrument of the Policy Experimentation Questionnaire of the University of Potsdam, at the beginning of the workshop on Friday evening and at the end of the workshop on Sunday afternoon, as well as with focus group sessions in which the participants were interviewed by the instructors and assistants according to a questionnaire from WP4.

Friday, 24.11.2023

Before 18:00	Participants arrive and check into their rooms
18:00	Dinner
19:00 - 21:30	Getting to know each other Policy Experimentation Questionnaire (I) Introduction to AI (act_02) Activity: generating cocktail recipes with AI (act_28)
21:30	Finish the evening: mixing fruit cocktails

Saturday, 25.11.2023

8:00	Breakfast
9:00 - 10:45	Introduction to Data literacy and the data lifecycle (act_06) Abalone: multiple linear regression with orange3 (act_25)
11:00 - 12:30	“Creativity and Data: Food pairing” - presentation by an expert (act_12)
12:30	Lunch
13:30 - 14:00	Kickoff data workshop “Best Pizza in town” (act_29)
14:00 - 15:30	City walk Zeitz / Data collection
15:45 - 19:00	Data workshop (act_29)
19:00	Pizza - topped by the participants and baked in a stone oven
21:00	Open Space: (act_13) Prompt Battle etc. (act_30)

Sunday, 26.11.2023

8:00	Breakfast
9:00 - 11:30	Review data workshop: current status, results and achievements Learning stations: ● Data workshop Price-Predictor-App (act_29) ● Food quality control and algorithmic bias (act_27)

	• Grimm's new tales - Language models unplugged (act_16) • Decision trees in Orange (act_07) • Teachable Machine (act_19, act_24)
11:30-12:30	Discussion: Data competences for the creation of AI applications
12:30	Lunch
13:30 - 15:00	Evaluation, Closing Policy Experimentation Questionnaire (II) Focus group sessions
15:00	Departure of the participants

Implementation

As in 2022, we announced the event via various mailing lists of the German Informatics Society (GI) and teachers' associations. We also wrote to teachers who had already taken part in our training courses. We received a total of 25 applications. We accepted all interested parties.

With one cancellation due to illness, 24 teachers traveled to Zeitz. Due to further illnesses, three participants unfortunately had to return home early on Saturday, meaning that we ended the workshop with 21 participants.

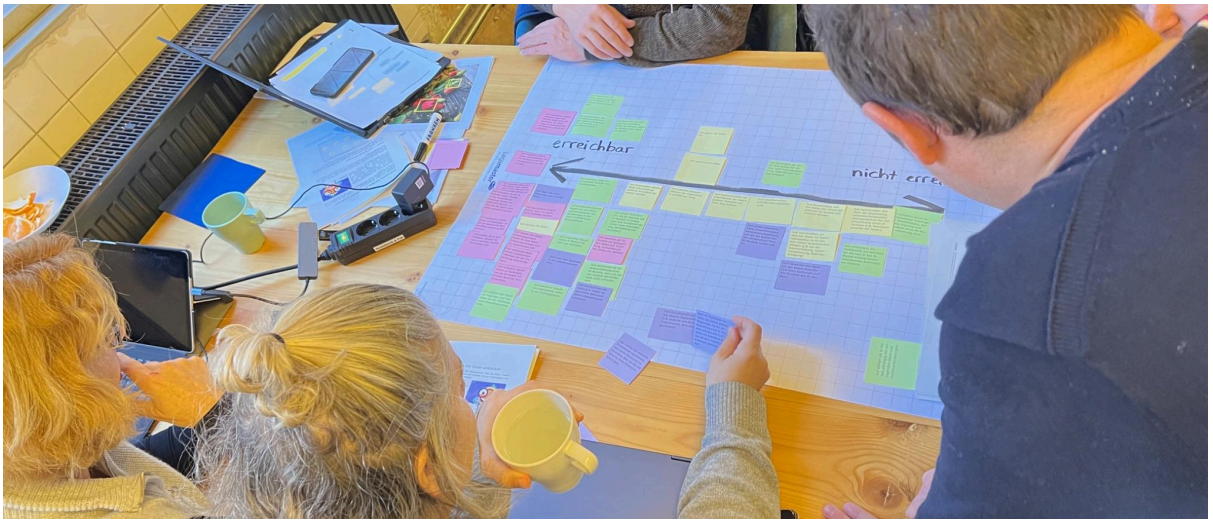


Figure 29: participants discussing data competencies as prerequisites to understand and develop AI applications

The workshop went largely as planned with a few deviations. The following are some of the instructors' impressions. A full report can also be found on the TrainDL website.[20]

- The participants had difficulties with the "Generative Cocktail Producer" activity. It was not clear in all steps of the instructions what was expected, or the built-in degrees of freedom for decisions (e.g. recording different quantities) proved to be too great. In addition to the built-in difficulty (a recipe for a Margherita pizza instead of a Margarita), there was a need for further adjustments. While the collection of

[20] https://train-dl.eu/en/news/detail/bits-bytes-and-tidbits-a-traindl-weekend-workshop

data and the creation of a knowledge base was still successfully implemented, only around half of the groups were able to generate new cocktail recipes.

- The data workshop included a visit to a few pizza restaurants to record their menus for the data project. One of the lessons learned from the data workshop is that a data project has similar characteristics to a software project, such as an open task, a variety of solutions that are not planned down to the last detail, the self-organization of the learners, the need for situational decisions and the possibility of failure.
- The social highlight of the workshop was baking pizzas together on Saturday evening. For a workshop lasting several days to succeed, it is important to give participants the opportunity to get to know each other and network.
- The Open Space on Saturday evening was used by one of the participants to present ENTER[21], an educational tool that illustrates the decision tree learning algorithm. We also played a few rounds of prompt battle, where two people or groups compete against each other to generate suitable images for a given task using an AI tool.
- A Miro Board, which we set up to accompany the workshop. The board reflected the structure of the workshop. Results were presented here and resources were made available (see Figure 30).
- The teachers liked the tactile approach of the Teachable Machine exercise, meaning that the students work with real objects and create images from them. However, they pointed out that only a few schools have access to PCs with cameras for use in teaching. Also, some participants experienced difficulties when using the built-in cameras from their laptops, as they had to hold the objects up to the camera which led to unclear images or images in which the hand was more prominent than the object. In the current version, although it is a web application, Teachable Machine unfortunately does not work on iPads or smartphones.There were technical issues with the framework in an optional part of the exercise where the learners integrate image recognition into a Scratch program.

[21] https://d3rfuch5.github.io/ENTER/index.html

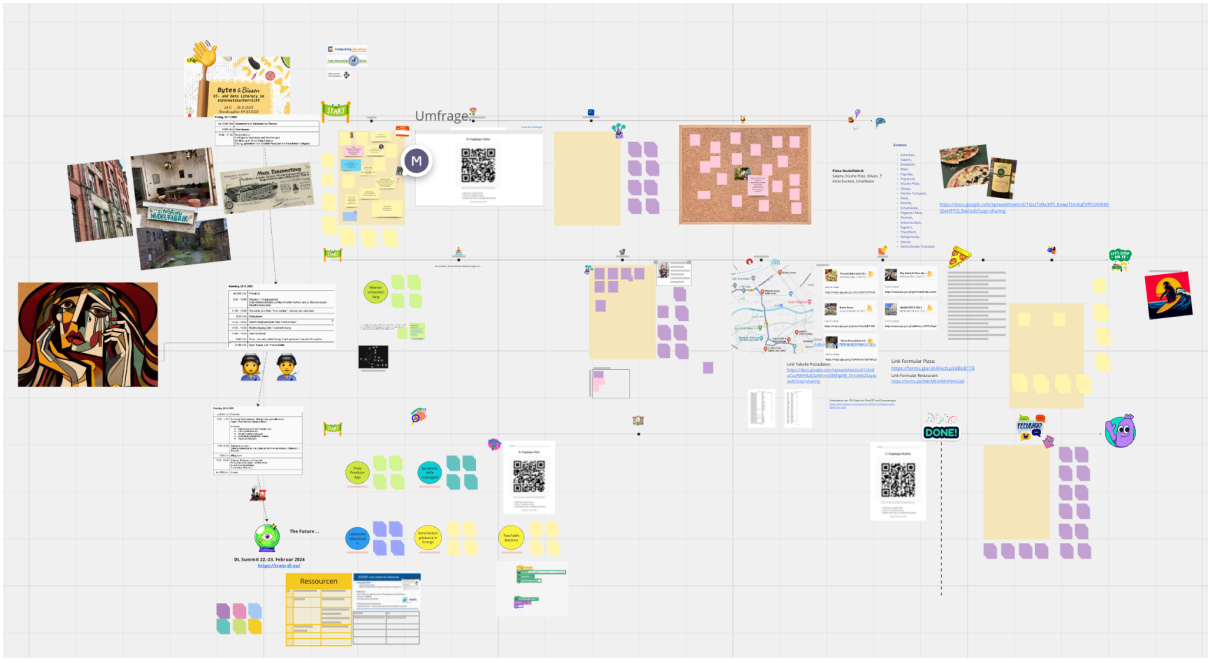


Figure 30: An overview of the Miro Board that accompanied the event.

09.01.2024 | Vilnius | Secondary school teachers

Figure 31: Participants in the STEAM teacher training in Vilnius, 09.01.2024

15 teachers took part in the teacher intervention in Vilnius in January. Teachers subjects were physics, foreign languages, art, technology, informatics, mathematics, primary education teachers. The course of the workshop was the same as the training for secondary school teachers in Graz on 23.10.2023. During the intervention, the theoretical part of artificial intelligence and data literacy was reviewed, as well as monkey grouping and food quality analysis tasks. The teachers enjoyed the food quality analysis task very much and the task itself showed very well how the choice of initial data can determine the final results. Teachers put it into practice and at least a few of them asked for the task material to use in their lessons. The teachers were not able to fully complete the abalone task, but after sorting the food. However, the teachers understood the principles and the purpose of the task itself. Some of the teachers mentioned that they already use ChatGPT in their lessons, and also expressed the wish that such workshops would not be a one-off, but continuous and it would be great to receive materials. The photos below show some of the moments from the workshop.

26.01.2024 - 28.01.2024 | Meißen | Secondary school teachers

The last of our three weekend workshops was aimed at secondary school teachers who do not teach computer science. We could assume that the teachers were interested in AI and data literacy, as they registered for the training programme voluntarily.

Figure 32: Logo for the weekend workshop in Meißen created by DALL-E

Schedule

The schedule was based on the tried and tested scheme for weekend workshops in Heidelberg and Meißen. In terms of content, we focussed on the heterogeneous target group of teachers from all subjects. However, the different content and teaching methods depending on the school subject made it difficult for us to simulate teaching situations and put the participants in the role of pupils (the pedagogical approach of double-decker that we used in all prior interventions). To enable the participants to focus on the specific needs of their subjects and exchange subject-specific information, we formed groups of teachers from the same subjects. In these groups, the teachers worked together several times throughout the workshop and finally developed lesson outlines for their subject, with or about AI and data literacy.

Following the Dagstuhl Triangle model, teaching technical perspective and socio-cultural perspective was based on the content that we taught in our previous training. We used some of our tried-and-tested activities from previous interventions to give participants an introduction to AI (Grimm's new tales, AI quiz). On Saturday morning we gave an introduction to the data lifecycle and let the participants work on our activity on multiple

linear regression. This allowed participants to learn about data analysis and experience how a non-computer science lesson applying machine learning as a scientific method could be designed. To cover the topic of AI ethics, we used the food quality control activity on algorithmic bias.

In order to make the applicational perspective of AI and data literacy understandable for the participants, we recruited two experts for the workshop to share their expertise. Marco Block-Berlitz[22] uses AI methods and quadcopters for 3D mapping of historical sites. Youssef Nader[23], a researcher from Freie Universität Berlin, is one of the winners of the Vesuvius Challenge[24] for deciphering charred papyri using AI methods. In addition to their presentations, the experts were to be available on a panel to answer participants' questions, even after their presentations. In this way, we were able to illustrate how diverse AI processes can be used even in applications, and also impart content knowledge to the participants.

Friday, 26.01.2024

Before 18:00	Participants arrive and check into their rooms
18:00	Dinner
19:00 - 21:30	Getting to know each other Formation of groups: Teachers of similar subjects work together Policy Experimentation Questionnaire (I) Grimm's new Tales: language models unplugged (act_16)
21:30	End the evening in the cafeteria

Saturday, 27.01.2024

8:00	Breakfast
9:00 - 9:30	AI Quiz (act_20)
9:30 - 10:45	Introduction to the Data Lifecycle (act_06) Abalone: predict the age of sea-snails with linear regression (act_25)
11:00 - 12:30	"Methods of artificial intelligence in archaeology" Presentation by Marco Block-Berlitz (act_12)
12:30	Lunch
13:30 - 15:00	Guided tour Meißen
15:15 - 16:00	"Resurrecting Ancient Scrolls with AI" Presentation by Youssef Nader (act_12)

[22] http://www.marco-block.de/
[23] https://youssefnader.com/
[24] https://scrollprize.org/

16:00 - 16:45	“What you always wanted to know about AI: Teachers ask, AI experts answer” Expert panel (act_12)
17:00 - 18:30	AI Chatbots at school (act_22)
18:30	Dinner
19:30	Food quality control: algorithmic bias (act_27)
20:30	Open space (act_13)

Sunday, 28.01.2024

8:00	Breakfast
9:00 - 10:30	AI-PACK: "AI and data: What skills do teachers need?" (act_31)
10:45-12:00	Group work: Development of a teaching project with or about AI and data literacy (act_14)
12:00	Lunch
13:00 - 14:30	Presentation of teaching projects Evaluation, Closing Policy Experimentation Questionnaire (II) Focus group sessions
14:30	Departure of the participants

On Sunday morning, we introduced the AI-PACK competence model and let the teachers reflect on the AI competencies for teachers.[25] Finally, the emphasis was on how participants could take the content from the workshop and incorporate it into their own teaching practice through the development of the teaching project mentioned above.

Implementation

Advertising the workshop proved to be a challenge, as we could not rely on our professional network of CS teachers as we did for the previous weekend workshops. 23 participants registered for the workshop. It was noticeable that many teachers held management positions at their schools or were involved in the training of trainee teachers. Following subjects were represented during the workshop:.

- Music
- Languages
- Math & Physics
- Social sciences
- Biology & Chemistry

[25] https://dl.gi.de/items/6bb42c88-f06b-4b9b-accd-47da7b8d7ed6

The fact that a rail strike was called a few days before the workshop forced participants to reschedule their journeys. We set up an online message board to encourage the formation of car pools, which worked out excellent.

Observations made by the instructors during the workshop:

- The teachers decided to form a unified subject group comprising Math, Physics, Biology and Chemistry. This resulted in one big group of more than ten people. At the beginning we encouraged them to split the group, but they enjoyed working in a group of that size.
- When working with the AI quiz, the participants showed a great interest in discussing definitions and applications of AI. We had not anticipated this. To make room for this need, we extended the time for the AI quiz and moved the linear regression activity to Saturday night. We attribute the need for discussion to the fact that we had a field with different school subjects and correspondingly diverse perspectives on artificial intelligence. For a future workshop, we could expand the AI quiz and confront the groups with different definitions of AI or have them develop their own.
- The expert presentations were highly praised and considered by many to be the highlight of the workshop.

Figure 33: Youssef Nader presents his approach to decipher charred papyri scrolls with the help of machine learning and to solve the Vesuvius challenge.

- Several participants contributed questions to the expert panel. However, it would have been a good idea to include a preparatory phase in which the participants could discuss the content of the presentations and formulate questions.
- We received feedback that Saturday was perceived as very exhausting by some participants because there was a lot of input. Also we were still working after dinner while in Zeitz we had spent the evening preparing pizza and having an open

space afterwards. Consequently, the open space was not used by any participant to present their own work.

- On Sunday, one participant shared his reflections on the use of AI for lesson preparation as well as for teaching social science for grades 5-6.[26]
- After the successful use of the Miro board in the Zeitz intervention we continued the good practice and set up a board for the Meißen intervention as well (see Figure 34).

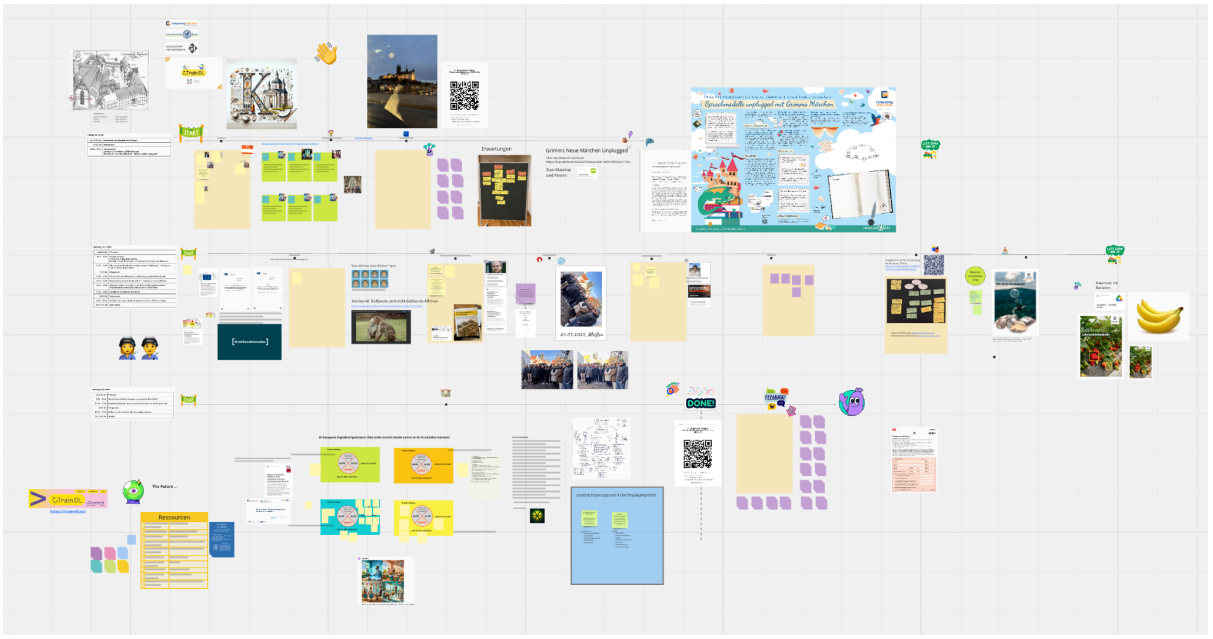


Figure 34: Overview of the collaborative Miro board with supporting materials and work from the participants. The view is zoomed out to protect the participants' private data.

[26] https://www.mieriesuperklasse.de/KI_in_gewi/index.html

Challenges

In this chapter, we summarize the most important challenges that needed to be addressed during the third round of interventions.

AI Hype

With the release of ChatGPT in November 2022, a big milestone in the development of intelligent chatbots since ELIZA (1966), artificial intelligence became a hype topic in society's mainstream. On one hand, this led to a great interest in the topic, also in the educational field, and brought the TrainDL project a lot of attention. We were faced with the challenge of quickly conceptualizing the new developments in the field of generative AI, placing them in the educational system environment and integrating them into our workshop concept.

Our first attempt in this regard was made in March 2023 in the second round of interventions with a presentation focused on the school context and an unplugged activity on language models (Grimm's new fairy tales). The challenge for the third round of interventions was to present the topic suitable for STEAM teachers. On the one hand, we also tested the use of the Grimm activity, which was constantly developed during the third round of interventions. On the other hand, we took into account the fact that in the professional everyday life of non-computer science teachers, the use of generative AI tools plays a bigger role than its functional understanding. Therefore, we developed an activity in which teachers can exchange ideas about their use of AI chatbots, try out prompts, and develop teaching concepts.

Workshop concept for STEAM teachers

Compared to professional development for computer science teachers, creating a workshop concept for STEAM and then all secondary school teachers was challenging. Firstly, there was an internal discussion in the project about how to differentiate between computer science skills and media literacy in the field of artificial intelligence and data literacy. As described in our project proposal, the focus of the TrainDL project was on classroom-based professional development. A prompt training, like those offered by various commercial providers, where teachers learn, for example, how to write a letter to the parents with a chatbot or generate a lesson plan, would have missed the project's goals.

One learning from the intervention in Düsseldorf (August 2023) was that non-computer science teachers have little interest in understanding technological knowledge of AI and data literacy such as the paradigms of machine learning or the data lifecycle. Though a subject-specific approach, as we pursued in the activities on linear regression using the example of sea snails or in the station on musical AI applications, was well received, we still had two challenges. On the one hand, the effort to develop suitable activities for specific subjects is high and adopting the needs of all subjects in one activity is not possible. On the other hand, as computer science educators, we do not have subject-specific expertise and rely on collaborations with other subject didactics for material development that meets scientific standards.

Future research should investigate more deeply which AI and data literacy skills teachers of non-computer science subjects need. This field is in rapid development through the pace of technological progress. Therefore, future training sessions should focus all the more on fundamental ideas and not on fleeting technological developments. One promising approach is to follow the AI-PACK model, which we have introduced to the teachers in the weekend workshop in Meißen.

List of Abbreviations

AI	Artificial intelligence
CK	Content Knowledge
CS	Computer science
DL	Data Literacy
LLM	Large Language Model
ML	Machine Learning
OCG	Austrian Computer Society
STEAM	Science, Technology, Engineering, the Arts, and Math
TrainDL	Train Data Literacy
UP	University of Potsdam
WP2	Work package 2
WP4	Work package 4

Appendix

Ankündigung STEAM workshop in Berlin

Titel: KI und Datenanalyse im MINT-Unterricht
Zielgruppe: Lehrerinnen und Lehrer aller MINT-Fächer
Lernziele:

- KI-gestützte Datenanalyse als neues Unterrichtswerkzeug kennenlernen.
- Problemstellungen aus den MINT-Fächern mit dem Tool “Orange3” bearbeiten.

Dauer: 90 Minuten

Beschreibung:
Der Workshop 'KI und Datenanalyse im MINT-Unterricht' bietet Ihnen eine praxisnahe Einführung in die KI-gestützte Datenanalyse, die in einer Zeit, in der Künstliche Intelligenz immer mehr Anwendungen findet, für Schülerinnen und Schüler eine Schlüsselkompetenz darstellt. Durch den Einsatz von KI-Werkzeugen im MINT-Unterricht können Schülerinnen und Schüler neue, innovative Herangehensweisen an Probleme erlernen.

Anhand von konkreten Fallbeispielen (Case Studies) und ausgewählten Datensätzen lernen Sie, Probleme mit Hilfe von KI-Werkzeugen zu lösen. Die Vorgehensweise orientiert sich dabei am Daten-Lebenszyklus und beinhaltet die Exploration, Analyse und Visualisierung von Daten. Im Schulunterricht ermöglicht der datengetriebene Lösungsansatz den Schülerinnen und Schülern neue Perspektiven auf Fragestellungen der MINT-Fächer.

Bitte installieren Sie das Datenanalyse-Werkzeug Orange3 [1] und bringen Sie Ihren Laptop zum Workshop mit. Programmierkenntnisse sind nicht erforderlich.

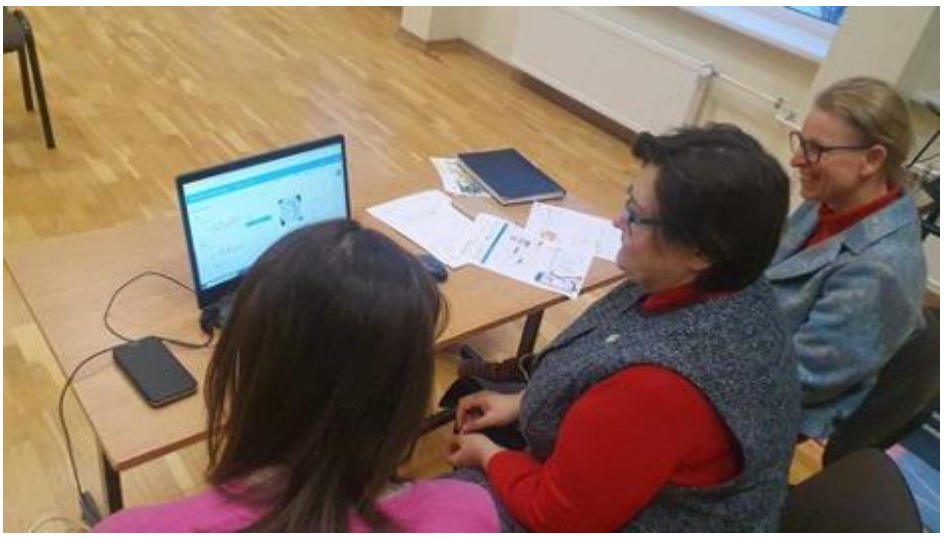
[1] https://orangedatamining.com/

Zur Workshopbeschreibung auf der MNU-Seite:
https://www.mnu.de/extern/mitglieder/anmeldung/teilnehmer/DetailsBeitrag.php?Beitrag=4876

- Veranstaltungs-Website: https://mnu.de/lv/bb/index.php/fortbildungen/landestagung
- Kontakt: XX
- FU Berlin, Silberlaube, KL29/208
- 11:30 - 13:00

Foto documentation

09.01.2024 | Vilnius | STEAM teachers



26.01. - 28.01.2024 | Meißen | STEAM teachers

ComputingEducation

Freie Universität Berlin

An die Lehrkräfte der natur-, gesellschafts- und geisteswissenschaftlichen Fächer

Institut für Informatik
AG Didaktik der Informatik
Königin-Luise-Str. 24/26
14195 Berlin

Einladung zur Fortbildung
"Künstliche Intelligenz und Daten im Unterricht - Mehr als nur ChatGPT?"

Liebe Kolleginnen und Kollegen,

Daten und Maschinelles Lernen sind ein fester Bestandteil des Alltags von Schülerinnen und Schülern. Dabei sind junge Menschen nicht nur diejenigen, die auf sozialen Medien online sind und mit intelligenten Algorithmen ihre persönlichen Daten teilen, sondern auch diejenigen, die KI-Technologie aktiv und bewusst nutzen, um kreativ zu sein oder zu spielen. **KI und Big Data verändern alle Disziplinen**, ob Natur-, Gesellschafts- oder Geisteswissenschaften. Entsprechend hoch ist der Stellenwert dieses Themas für den Unterricht. Seit Jahren stellen wir uns in der Informatikdidaktik folgende Fragen:

> Wie verändert KI die Unterrichtsfächer?
> Welche neuen Unterrichtsmethoden werden durch KI möglich?
> Was ist ein Prompt und was bedeutet Data Literacy?
> Wie rüsten wir unsere Schülerinnen und Schüler für eine von Künstlicher Intelligenz und Daten geprägte Zukunft?

Im Rahmen eines **Wochenendworkshops vom 26.-28.1.2024** möchten wir unsere vorläufigen Antworten darauf mit Ihnen teilen und diskutieren. Dafür laden wir Sie herzlich in den Klosterhof St. Afra nach **Meißen** ein. Die Fortbildung richtet sich vorwiegend an Lehrkräfte außerhalb der Informatik, die an einer weiterführenden Schule unterrichten.

Nach einer thematischen Einführung lernen Sie Werkzeuge wie Orange3 oder CODAP kennen und führen eigene Datenanalysen durch. Ein Vortrag des eingeladenen Daten-Experten wird uns zeigen, wie gravierend KI die Welt verändert und worauf wir unsere Schülerinnen und Schüler vorbereiten sollen. Sie erhalten Einblicke in die Funktionsweise von Chat-KIs und üben die Erstellung von Prompts. Im Team mit anderen Lehrkräften Ihres Faches oder verwandter Disziplinen entwickeln Sie im Workshop spezifische Unterrichtsprojekte. **Ziel des Workshops ist es,** Ihr fachdidaktisches Repertoire mit den Methoden der Datenanalyse und Künstlicher Intelligenz zu erweitern. Auch neben dem offiziellen Programm wird es reichlich Gelegenheit zum Austausch geben, etwa auf einer Tour durch die Meißener Altstadt.

Haben wir Ihr Interesse geweckt? Dann **schreiben Sie bitte bis zum 26.11.2023 eine E-Mail** mit dem Betreff "Meißen24" an zoppke@inf.fu-berlin.de mit den folgenden Angaben:

> Was motiviert Sie für die Teilnahme an der Fortbildung?
> Welche Erfahrungen mit KI und Daten im Unterricht haben Sie bereits gemacht?
> Welche Fächer und an welcher Schule unterrichten Sie?

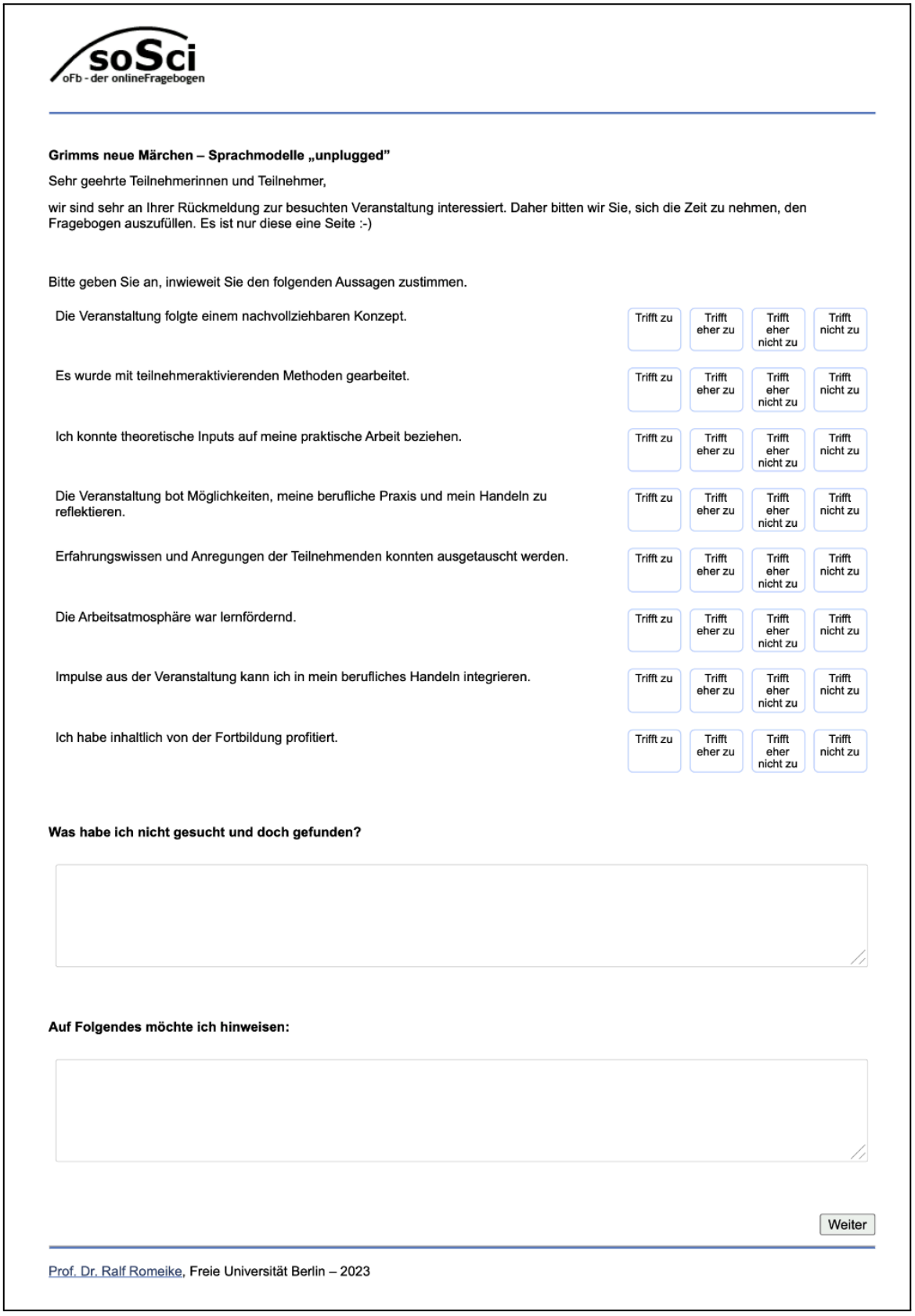
Wir können bis zu 20 Lehrkräften die Teilnahme am Workshop ermöglichen und geben Ihnen kurz nach dem Stichtag Bescheid, ob Sie einen Platz erhalten. Alle Kosten (Unterkunft, Verpflegung, DB-Fahrkarten) werden von der Freien Universität Berlin übernommen.

Prof. Dr. Ralf Romeike, Viktoriya Olari und Till Zoppke
07. November 2023

Das Wichtigste in Kürze
Ort: Klosterhof St. Afra, Freiheit 16, 01662 Meißen
Zeit: Freitag, 26.1. (18:00 Uhr) bis zum 28.1.2024 (15:00 Uhr)
Kostenübernahme: DB-Fahrkarten, Unterkunft und Verpflegung
Bis zu 20 Lehrkräfte
Anmeldung bis zum 26.11.2023

Invitation to the Meißen workshop

18.10.2023 | Berlin | STEAM teachers



soSci
oFb - der onlineFragebogen

Grimms neue Märchen – Sprachmodelle „unplugged"

Sehr geehrte Teilnehmerinnen und Teilnehmer,

wir sind sehr an Ihrer Rückmeldung zur besuchten Veranstaltung interessiert. Daher bitten wir Sie, sich die Zeit zu nehmen, den Fragebogen auszufüllen. Es ist nur diese eine Seite :-)

Bitte geben Sie an, inwieweit Sie den folgenden Aussagen zustimmen.

Die Veranstaltung folgte einem nachvollziehbaren Konzept.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu
Es wurde mit teilnehmeraktivierenden Methoden gearbeitet.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu
Ich konnte theoretische Inputs auf meine praktische Arbeit beziehen.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu
Die Veranstaltung bot Möglichkeiten, meine berufliche Praxis und mein Handeln zu reflektieren.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu
Erfahrungswissen und Anregungen der Teilnehmenden konnten ausgetauscht werden.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu
Die Arbeitsatmosphäre war lernfördernd.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu
Impulse aus der Veranstaltung kann ich in mein berufliches Handeln integrieren.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu
Ich habe inhaltlich von der Fortbildung profitiert.	Trifft zu	Trifft eher zu	Trifft eher nicht zu	Trifft nicht zu

Was habe ich nicht gesucht und doch gefunden?

Auf Folgendes möchte ich hinweisen:

Weiter

Prof. Dr. Ralf Romeike, Freie Universität Berlin – 2023

Online-questionnaire implementing the paper-survey for the workshop on 18.10.2023