AI4T
AI FOR TEACHERS

Work Package 3:

Evaluation

Deliverable D3.3

Comparative European

Evaluation Report

Co-funded by the
Erasmus+ Programme
of the European Union

AI4T
AI FOR TEACHERS

Project Title:	Artificial Intelligence for and by Teachers
Project Acronym:	AI4T
Project Number	626154
Grant Agreement Number	626154-EPP-1-2020-2-FR-EPPKA3-PI-POLICY
Deliverable number	D3.3
Work Package	3 Evaluation
Work Package Leader	Centre national d'étude des systèmes scolaires – Conservatoire national des arts et métiers (Cnesco-Cnam)
Work Package Partners	Dublin City University (DCU), Istituto Nazionale di Documentazione, per l'Innovazione e la Ricerca educativa (INDIRE), Pedagoški inštitut, Université du Luxembourg
Dissemination level	public
Delivery date	2023-01
Status	Final
Version	1
Author(s)	Aurélie Paris, Aude Labetoulle
List of contributor(s)	Simona Bezjak, Deirdre Butler, Pedro Cardoso-Leite, Jean-François Chesné, Christiane Kirsch, Paola Nencioni, Jessica Niewint, Lina Rivera, Francesca Rossi, Francesca Storai, Plamen Vladkov Mirazchiyski, Sara Mori
Deliverable Manager	Cnesco-Cnam, France Education International

Co-funded by the Erasmus+ Programme of the European Union

ABSTRACT	This report presents and compares the impact of the AI4T– Artificial Intelligence for and by Teachers – professional learning pathways in France, Ireland, Italy, Luxembourg and Slovenia. The first sections introduce the AI4T professional learning pathway and outline the experimental design in each country, i.e. participant selection and randomisation procedures, theoretical framework used for assessment, and evaluation instruments. This is followed by a description of the country samples. Finally, issues pertaining to the experiment's internal and external validity are addressed. A comparative perspective is adopted to present the findings across the different countries. The primary focus is given to teachers as they are the main target of the AI4T project. After detailing their reactions to the professional learning pathways in the five countries, the report discusses the three main outcomes of the experiment, namely teachers' knowledge of AI, their perceptions of AI, and their use of AI. Both the initial state and the impact of the intervention are presented for each outcome. In addition, we provide a discussion of the differential impact of the intervention on teachers based on their engagement with the MOOC, their self-efficacy for integrating technology in the classroom, and their course subjects. Results from school leaders and students are then presented in separate sections. A final section highlights common takeaways from all five countries which are likely to inform future AI-based educational policies. The section focuses on specific needs in terms of professional learning, tool development and ethical safeguards.
KEYWORDS	Artificial intelligence, experimentation, evaluation, impact study, professional learning, teachers

Dissemination level		
PU	Public	**X**
PP	Restricted to project partner (including the Commission)	
RE	Restricted to a group defined by the consortium (including the Commission)	
CO	Confidential, only for members of the consortium (including the Commission)	

Co-funded by the Erasmus+ Programme of the European Union

AI4T
AI FOR TEACHERS

Table of content

Co-funded by the
Erasmus+ Programme
of the European Union

List of figures

List of tables

Co-funded by the Erasmus+ Programme of the European Union

Co-funded by the
Erasmus+ Programme
of the European Union

Introduction

In recent years, the rapid development of new technologies based on Artificial Intelligence (AI) has prompted a crucial discussion on its implications for education. At the European level, the *Digital Education Action Plan 2021-2027* emphasises the need to develop students' AI skills and to provide ethical guidelines on the topic.

Funded by the European Commission, the *Artificial Intelligence for and by Teachers (AI4T)* project was a three-year experiment designed to explore and support the use of AI in education. It consisted in producing, implementing and evaluating professional learning activities to familiarise teachers with AI. The project was conducted in 5 countries: France, Slovenia, Italy, Ireland and Luxembourg. 17 partners, including education ministries, evaluators and research centres took part in the project, under the coordination of *France Education International (FEI)*.

The AI4T intervention was built around two common online resources developed for the project and common to the five countries: the AI4T MOOC created under the coordination of the *Institut national de recherche en sciences et technologies du numérique* (Inria) and the *AI for Teachers: An Open Textbook* manual written under the coordination of the *Université de Nantes*. Both resources received contributions from the consortium partners. In each country, professional learning pathways were developed, based on common learning objectives but using a variety of formats (e.g. online platforms, webinars, face-to-face sessions).

Following a pilot study conducted in 2021-2022 with a small sample of schools, the intervention was implemented in the 2022-2023 school year. The study focused on maths and language teachers with students aged 15 to 17. A few teachers from other subjects such as experimental science or computer science were also included in the experiment. Within each country, except Luxembourg, half of all participating schools were randomly chosen to initially participate in the intervention. In this random school sample, teachers attended the professional learning pathway during the experimentation year. Teachers in the remaining schools constituted a control group and were given access to the AI4T resources after the end of the experiment only.

Data collection methods included surveys administered to teachers, school leaders, and students, and interviews conducted with teachers and school leaders. These findings are complemented with elements from the analysis conducted by a partner of the project, the *Laboratoire lorrain de recherche en informatique et ses applications* (Loria) on teachers' learning traces[1] on the MOOC. Based on the data collected, this report addresses the following four evaluation questions:

1) *Was the professional learning experience conducive to <u>teachers' learning of AI</u>?*
2) *Was the professional learning experience conducive to changing <u>teachers' perceptions of AI</u>?*
3) *Was the professional learning experience conducive to modifying <u>teachers' use or behavioural intentions of using AI</u>?*
4) *What are some key factors that can account for the impact of the intervention?*

[1] The learning traces correspond to the digital traces left by users on the MOOC.

Co-funded by the Erasmus+ Programme of the European Union

1. Intervention

The AI4T intervention revolved around two common online resources, both translated for all 5 participating countries. The first resource is the AI4T MOOC created under the coordination of the Inria. The second resource is a textbook entitled *AI for Teachers: An Open Textbook,* produced under the coordination of the *Université de Nantes*. It was offered as an additional resource for more experienced users and instructors. The following outcomes underpinned the professional learning pathways for the teachers in all 5 countries:

1. Being able to express one's understanding and attitude towards AI and discuss it.
2. Being able to understand the basic principles of AI systems.
3. Being aware of AI educational applications and key considerations when identifying, assessing and selecting an AI tool for teaching, learning and assessment.
4. Being aware of legal considerations when using AI in an educational setting.
5. Being aware of ethical considerations when using AI in an educational setting.
6. Being aware of generic AI tools and being able to reflect on their impact on education and critically consider the possibilities for AI tools in education.

Based on these common resources and learning outcomes, professional learning pathways, which incorporated complementary sessions, were designed in each country.

The **format, sequencing and content of these complementary sessions** differed from country to country. Some countries included a mix of face-to-face sessions and online sessions (France, Ireland, Luxembourg), while others decided to use online sessions only (Italy, Slovenia). The **sequencing and interconnectedness** of the MOOC sessions and of the learning events varied. In Ireland, Italy and Luxembourg, for instance, the pace of study on the MOOC was guided and periods of self-study on the MOOC were followed by online group sessions; whereas, in France and Slovenia, teachers could work at their own pace on the MOOC and the other learning events were stand-alone events. A feature of the Italian pathway was that the intervention group was divided into smaller groups. Each of these sub-groups was assigned a tutor, and every session was consistently led by the same tutor.

In some countries, representatives were available to answer questions using dedicated platforms. For example, in France, representatives answered participants' questions on the MOOC's forum.

The manner in which the **textbook** was integrated also varied from country to country. In Luxembourg, France and Slovenia, the textbook was introduced during a face-to-face session or a webinar, but at a different stage of the intervention (e.g. a few days before the endline questionnaire in France). In Ireland, the textbook was used as a tool by the professional learning advisors to facilitate the online sessions and it was promoted to the learners only after the intervention.

Although all countries presented **AI tools** to teachers, these tools varied from one country to another: ChatGPT, MidJourney and Vittascience in France, InstaText and Orange in Slovenia, AI in Microsoft Office applications, ChatGPT, DallE and Duolingo in Luxembourg, ChatGPT, Duolingo and Photomath in Ireland, and a variety of AI-based language tool (e.g. Duolingo, Deepl), AI-based STEM tools (e.g. Cymath, MathSolver) and generative AI tools (e.g. Dall-E, ChatGPT) in Italy.

A distinctive aspect of the French experiment was the provision of a licensed tool called Kwyk for mathematics teachers in both the intervention and control groups, allowing them to test the tool during the experimentation year. Kwyk is an online resource consisting of self-correcting maths exercises, corresponding to the French secondary school curriculum. Additionally, it provides teachers with analytical tools to better assess their students' difficulties.

Co-funded by the
Erasmus+ Programme
of the European Union

AI4T
AI FOR TEACHERS

Table 1: The professional learning pathways in the five countries

France	Ireland	Italy	Luxembourg	Slovenia
Online launch event: overview of the project (17/11/22) Online self-study: MOOC (04/01/23-25/03/23) Online experts webinar no.1: presentations on AI and learning analytics in education (25/01/23) Face-to-face session in 6 *académies*[2]: examples of AI tools, learning data, ChatGPT and generative AI (Feb – March 2023) Online experts webinar no.2: teaching with AI (15/03/23) Forthcoming: closing session (1st semester 2024) Amount of time dedicated to the professional learning pathway: 14h30, 3 months	Online launch event: overview of the project and introduction to AI (09/01/23) Online self-study: MOOC module 1; and online session to explore the contents of module 1 (19/01/23) Online self-study: MOOC modules 2 and 3; and online session to explore the contents of the modules with a focus on ChatGPT (07/02/23) Face-to-face, shared learning event to review the training pathway, debate and exploration of AI tools (01/03/23) Amount of time dedicated to the professional learning pathway: 15h[3], 2 months	Online launch event (24/02/23) General presentation of the MOOC (13-20/03/23) Webinar: introduction and socialization (20-21/03/23) Online self-study: MOOC module 1 (20-27/03/23) Webinar in small groups on generative AI (27-28/03/23) Online self-study: MOOC module 2 (27/03/23-3/04/23) Webinar in small groups on types of AI and AI tools for education. Online self-study: MOOC module 3, 4 and conclusion (3-24/04/23) Webinar in small groups on the use of AI tools in classrooms (20-21/04/23) Webinar in small groups on case studies (11/05/23) Amount of time dedicated to the professional learning pathway: 2 months	Face-to-face launch event: MOOC activities and baseline questionnaire (17/01/23) Online self-study: MOOC first chapters (17/01/24-01/02/23) Webinar: MOOC understanding and AI in Office 365 tools (01/02/23) Online self-study: MOOC last chapters (01/02/23-21/02/23) Face-to-face session: Content and activities in compliance with MOOC (21/02/23) Face-to-face session: Content and activities in compliance with MOOC (28/02/23) Webinar: Ethical guidelines for teachers (06/03/23) Face-to-face session: Presentation and discussion of task, closing, endline questionnaire (14/03/23) Amount of time dedicated to the professional learning pathway: 17h, 2 months	Online launch event: introduction to AI4T (13/02/23) Online self-study: MOOC (17/02/23-20/03/23) Webinar: presentation of the AI tool InstaText (22/02/23) Webinar: presentation of the textbook (28/02/23) Webinar: presentation of the AI tool Orange 1/2 (07/03/23) Webinar: presentation of the AI tool Orange 2/2 (20/03/23) Online closing meeting (20/03/23) Amount of time dedicated to the professional learning pathway: 24h45[4], 1 month and a half

[2] In France, "académie" refers to an educational unit, or local authority, that is based on geographical location.
[3] The number of hours presented for the Irish professional learning pathway does not include the time teachers engaged with the MOOC.
[4] The number of hours presented for the Slovenian professional learning pathway includes time spent by participants on the evaluation questionnaires.

Co-funded by the Erasmus+ Programme of the European Union

2. Experimental design

2.1 Participant selection and sample randomisation

The selection of teachers was carried out by the Education ministry in each country. The evaluation team submitted the same guidelines to the Education ministries. They were asked to select volunteer mathematics and modern languages teachers, working with students aged 15 to 17. The samples were not considered representative of the general population of teachers, but ministries were encouraged to select socio-economically diverse schools located in different areas. The guidelines for participant selection were adapted locally in response to recruitment difficulties. In Slovenia and Italy, selection was opened to teachers from other subjects, while selection in France was opened to teachers working with 14-year-old students. The recruitment period also differed from country to country, starting as early as May 2022 (France and Slovenia) and ending as late as March 2023 (Italy). Overall, 1003 teachers from 287 schools were selected to participate in the project. Once selected, each teacher identified the class with which they would run the experiment.

In all countries, except Luxembourg, the sample was randomly divided and assigned to one of two groups: an intervention group and a control group. While the intervention group was granted access to the AI4T professional learning pathway during the experimentation year, the control group was only granted access to the online learning resources after the end of the evaluation process. Participants were randomised at the school level. Following the recommendations of Banerjee and Duflo (2017), stratification was employed as the chosen method for randomisation. Stratification criteria were adapted from country to country depending on the sample size and on their relevance within the national context. Table 3 shows the criteria used for randomisation in each country. While some criteria were selected for logistical purposes, others were chosen because they had been identified as having a potential impact on attitudes to technology and technology integration. These criteria included sex (Poyet, 2015; Badia *et al*., 2014; Céci, 2019), course subject (Perotta, 2013) and the school's socio-economic make-up (Perotta, 2013). Participants were informed of their group – either the intervention or the control group – after the administration of the baseline questionnaire for teachers.

In countries with larger samples (France, Italy and Slovenia), a subset of schools from the intervention group was selected for the qualitative evaluation. Whereas in Ireland and Luxembourg, due to small sample size, all the participants in the intervention group were invited to take part in the interviews.

Co-funded by the
Erasmus+ Programme
of the European Union

Table 2: Participant selection in each country

France	*Ireland*	*Italy*	*Luxembourg*	*Slovenia*
Recruitment period: May-October 2022 Number of schools: 120 Number of teachers: 256 (142 maths teachers, 113 English language teachers)	Recruitment period: November 2022 Number of schools: 21 Number of teachers: 21 (12 maths teachers, 9 modern languages teachers)	Recruitment period: January-February 2023 Number of schools: 91 Number of teachers: 438 (262 STEM teachers, 152 English language teachers, 24 teachers of other subjects)	Recruitment period: October 2022- January 2023 Number of schools: 14 Number of teachers: 19 (12 modern languages teachers, 7 maths teachers)	Recruitment period: May-December 2022 Number of schools: 76 Number of teachers: 269 (121 maths teachers, 97 English language teachers, 51 teachers of other subjects)

Table 3: Stratification criteria for the randomisation in each country

France	Ireland	Italy	Slovenia
Region Type of school Index of the social composition of the school Number of volunteer teachers in the school Course subject Sex	DEIS status (indicator of the socio-economically disadvantaged level of the school) School size Course subject	Region Type of school Couse subject	Region Type of school Number of volunteer teachers

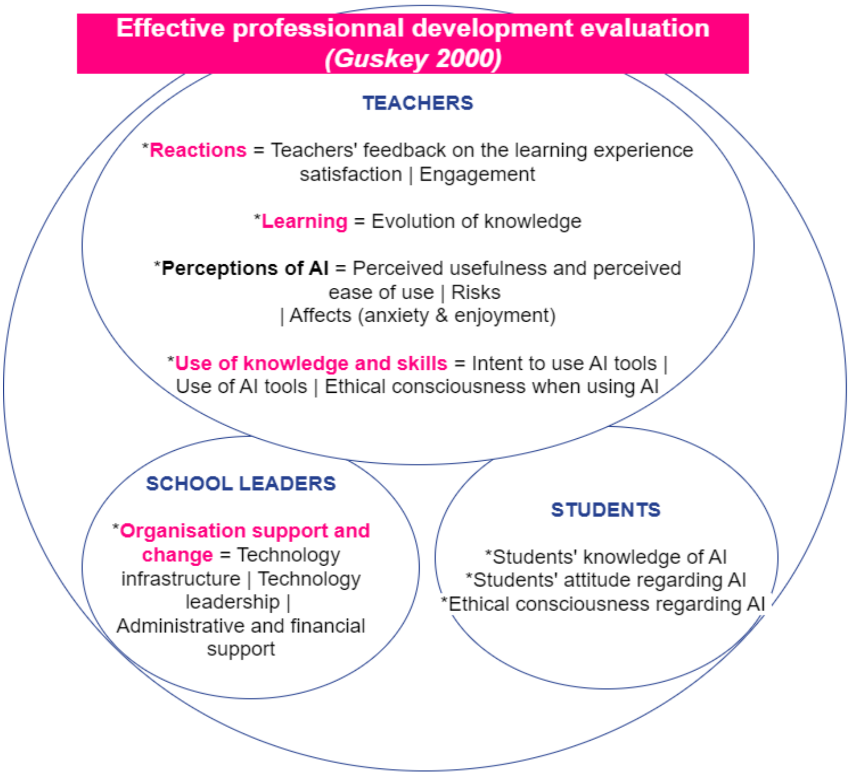
Co-funded by the Erasmus+ Programme of the European Union

2.2 Theoretical framework

AI4T started as a pioneer project on AI in education, tackling a relatively unexplored topic. To improve the evaluation questions asked at the beginning of the project, the evaluation team adopted a theoretical framework drawing from various domains of expertise, including AI but also digital technologies and professional development evaluation. Specifically, we drew from Guskey's work as a foundational framework (2000). According to Guskey, the effective evaluation of professional development requires the collection and analysis of five critical levels of information: (1) participants' reactions, (2) participants' learning, (3) organisational support and change, (4) participants' use of new knowledge and skills, and (5) students' learning outcomes.

For each level we created indicators adapted from existing scales and tested them during the pilot phase of the project. Scales were based on the Likert format and generally had 7 answer options for teachers and 5 for students. The response anchors were chosen following the recommendations of Casper *et al.* (2019) to ensure equal intervals between each anchor.

Figure 1: Theoretical framework for the evaluation of the AI4T professional learning pathway



Participants' reactions were assessed by measuring participants' engagement and satisfaction with the professional learning pathway. The **engagement** scale was adapted from Deng *et al.* (2020). The level of engagement with the professional learning pathway was measured using the behavioural, cognitive, social and emotional connections that the participants made with the course content, the instructors and with other learners. Behavioural engagement corresponds to learners' observable actions such as note-taking, while cognitive engagement corresponds to participants' cognitive investment in the learning process. Social engagement refers to both learner-instructor and learner-learner interactions, while emotional engagement centres on emotional connections with the professional learning pathway (enjoyment, interest, etc.). The **satisfaction** scale was adapted from Yenneck (2014). Yenneck identified key dimensions of satisfaction, such as satisfaction with the perceived usefulness of the course, which have an impact on learning benefits and ensuing changes in teaching practice. For both scales, participants were presented with statements and had to answer on

Co-funded by the
Erasmus+ Programme
of the European Union

a Likert scale from strongly disagree to strongly agree. The answers were then converted into scores from 1 to 7.

The measure of **participants' learning** was based on the content of the AI4T MOOC and additional reports on AI (European Commission, 2019; Samoili *et al*., 2020; Fengchun *et al.*, 2021). Experts on AI in education from and outside of the consortium were consulted to review the questions and their interpretation. To measure participants' learning, teachers were asked to assess their own knowledge of AI, indicate their level of familiarity with AI technologies, answer true/false questions on how AI works, and identify tools that contain AI. They were also asked through open questions to give a definition of AI and to name an AI tool that could be used for educational purposes.

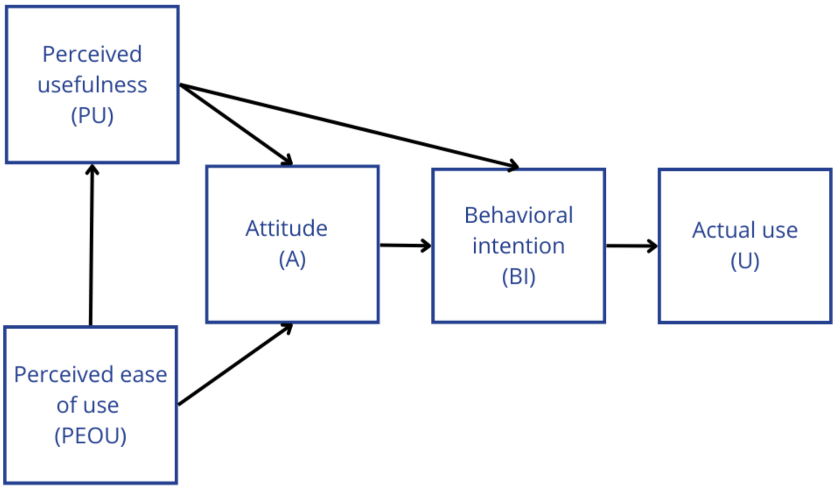
Data on **organisational support and change** were collected through school leaders. Guskey recommends assessing whether the organisation's policies and characteristics are compatible with the implementation of the changes being planned. To address the integration of AI, the evaluation team assessed the technological infrastructure and technology leadership of the schools. Access to technological equipment is sometimes described as the first-order barrier for technology integration, in comparison to the second-order barrier that is teachers' beliefs (Ertmer *et al*., 2012). A technological infrastructure is a prerequisite for integrating technology into teaching practices. The second variable to be measured, technology leadership, was developed by Anderson and Dexter (2005). In their model of technology leadership, Anderson and Dexter point to several indicators such as school leaders' own use of the technology. They stress the importance of school leaders setting an example by using the technology themselves in order to encourage its use in the whole school. Their indicators also include the number of days school leaders spend on planning, maintaining and administering the technology and the presence of an ethics policy within the school for the use of the technology. The evaluation team used these indicators to assess whether the school context was favourable to AI integration. Because Shattuck (2009) emphasises the importance of school leaders in upholding a vision for integrating technology that aligns with teachers' vision, we also included this factor in our measures. Finally, we assessed the administrative and financial support provided to teachers for their participation in the professional learning pathway.

Given the specific context of the project, which centres on changing teachers' **perceptions of AI** and encouraging the integration of AI tools in classrooms, an intermediate level was incorporated to the framework before **participants' use of knowledge and skills.** This level is based on the Technology Acceptance Model (Davis, 1989), described by Scherer *et al.* (2019, p.4) as follows:

> *In the literature, the question is repeatedly put forward as to what variables determine technology integration in education. Measuring user acceptance of technology is a way of determining the teacher's intentions toward using new technologies in their educational practice. Over the last decades, a series of models have been proposed to describe the mechanism behind and factors affecting technology adoption. […] Despite the variety of models, the TAM has dominated the research landscape as the most commonly used model to describe use intentions and actual technology use.*

Co-funded by the
Erasmus+ Programme
of the European Union

Figure 2: Technology Acceptance Model developed by Davis *et al*. (1989)



This model identifies two main variables, “perceived ease of use” and “perceived usefulness”, that determine behavioural intention to use and actual use of a technology. The evaluation team adapted the original scale from Davis *et al.* (1989) to measure “perceived ease of use of AI”. To measure the perceived usefulness of AI, we created items specific to the teaching profession, that enabled us to gain information on the specific teaching practices (identified by André Tricot, Cnesco, 2020) for which teachers perceived AI to be the most useful. In order to counter-balance the positive concept of “perceived usefulness”, we also studied participants’ answers on “risks” posed by AI, based on elements identified by Schiff (2021) and Remian (2019).

Some versions of the TAM also include the concept of “attitude”, whose definition and scope often varies (Njiku, 2019). We took a particular interest in one of the subdimensions of attitude, namely “affects”. Affects regarding AI are prominent in the AI literature (Wang and Wang, 2019, Cave *et al.,* 2019); they are also of interest to AI4T project partners, and they can further impact the use of technology (Février *et al.*, 2011). We therefore measured AI anxiety by adapting items from the Wang and Wang scale on AI anxiety (2019), and AI enjoyment by generating items based on existing scales on computer enjoyment (Christensen & Knezek, 2009; Noiwan *et al.*, 2005).

Both **behavioural intentions to use AI** and **actual use of AI** were measured, in accordance with the TAM. We also characterised the types of use by asking questions on the frequency of use, on the tools used, and on the tasks performed with the tools. Finally, we measured participants’ ethical consciousness when using AI by using items from a subscale on ethics in the AI literacy scale (Wang *et al.*, 2022).

Due to the characteristics of the AI4T professional learning pathway – objectives, duration and content – and the focus on teachers, we did not measure **student** learning outcomes, but we instead gathered contextual information on students’ knowledge, attitudes and ethical concerns regarding AI. We created an attitude-towards-AI-in-education scale based on the concept of attitude developed by Njiku *et al.* (2019) and on existing scales on attitude towards AI (Suh & Ahn, 2022; Shepman & Rodway, 2020). For the ethical concerns scale, we reviewed existing literature to include the main concerns mentioned in current research on AI in education (Jang *et al.*, 2022; Remian, 2019; Schiff, 2021; Akgun & Greenhow, 2021; European Commission, 2022; Holmes *et al.*, 2021).

Co-funded by the
Erasmus+ Programme
of the European Union

2.3 Evaluation instruments

The evaluation of the AI4T intervention is both quantitative and qualitative. Data was collected using questionnaires and interviews. The evaluation also draws on additional materials such as learning analytics provided by the teachers' activity on the MOOC and produced by the Loria. The evaluation protocol and instruments were submitted to an ethics committee in each country and approval was granted to all evaluators to conduct the study.

Teachers, their school leaders, and the classes of students selected by the teachers, were tasked with answering online questionnaires. To measure the impact of the AI4T professional learning pathway on teachers, teacher participants were asked to answer the questionnaire twice – first, at the beginning and then, at the end of the experiment – while school leaders and students answered the questionnaire only once, at the end of the experiment, for contextual information. Teachers and school leaders received generic links to the questionnaire. They were also given individual evaluation numbers for accessing the questionnaires. Students completed the questionnaire in class under the supervision of a school staff member. Students from the same class used their teacher's evaluation number.

The teacher questionnaires addressed the main outcomes: teachers' knowledge, perceptions and use of AI. In the baseline questionnaire, teachers were asked to provide some basic background information about themselves (e.g. sex, teaching experience, etc.). In the endline questionnaire, teachers who had participated in the intervention were asked questions relating to their engagement and satisfaction levels with the intervention. The school leader questionnaire was used to provide informational data on the general characteristics and technological infrastructure of the schools, and on the administrative and financial support available for teachers' professional learning and integration of AI in the classroom. Finally, the student surveys addressed students' understanding of AI, attitudes towards AI and ethical concerns regarding AI.

Interviews were conducted either with a subset of schools or with all the teachers in the intervention group, depending on the size of the sample in each country. Interviews took place after completion of the endline questionnaire in order to avoid creating a bias in participants who had either taken part in the interviews or not. Most interviews were conducted one-on-one, although in Italy interviews were conducted using focus groups.

The interviews focused on teachers' experiences with the professional learning activities and the AI tools. They addressed the factors already present in the questionnaires in order to provide a richer understanding of participants' answers. Teachers were also asked about their expectations and recommendations regarding AI policies.

Learning traces were collected by the Loria. They correspond to the digital traces left by users of the MOOC (i.e., teachers in the intervention group only). These traces were used to assess users' levels of engagement with the online materials (e.g., via the number of clicks or the consistency in watching video tutorials), and to identify types of learners through cluster analysis. In countries with a larger sample size (France, Slovenia and Italy), a correspondence table matched the IDs of the learning traces with the IDs entered in the survey. Thanks to this table, the evaluation team studied how engagement with the MOOC modulates the impact of the professional learning activities on teachers' knowledge, perceptions and use of AI. More information on learning traces and analytics can be found in the report *Methodological Framework For Data Collection and Learning Analysis* (deliverable D1.3).

Co-funded by the
Erasmus+ Programme
of the European Union

Table 4: Calendar of the evaluation of the AI4T project

	Pilot phase	Full scale								
Countries	**Sept 2021 - August 2022**	**Novembre 2022**	**Decembre 2022**	**January 2023**	**February 2023**	**March 2023**	**April 2023**	**May 2023**	**June 2023**	**July 2023**
France										
Italy										
Slovenia										
Ireland										
Luxembourg										

:	Pilot phase
	Baseline survey (teachers)
	Endline survey (teachers, school leaders and students)
	Professional learning for the intervention group
	Interviews (teachers, school leaders)

Co-funded by the
Erasmus+ Programme
of the European Union

3. Data

3.1 Sample characteristics

Altogether, the evaluation process garnered participation from 736 teachers, with the majority of teachers being from Italy, Slovenia, and France. 69.3% are female teachers – this proportion being lower in France (54.4%) and Luxembourg (30%), and higher in Italy (70.2%), Slovenia (79%) and Ireland (93%). Overall, teachers had a lot of teaching experience in all countries, with 46.9% having more than 20 years of teaching experience. Almost 80% of the sample was equally divided between maths and modern languages teachers. 11.4% were computer science or experimental science teachers (almost exclusively from Italy), and 9.2% taught other subjects. 50.5% of teachers came from academic schools, 32.9% from vocational schools and 12.2% from another type of school. The remaining 4.3% of teachers worked in French “collèges” corresponding to lower-secondary education. About half of the teachers had fewer than 25 students in their class. Important differences were observed among countries on this variable; for instance, in Italy and Luxembourg, 78.9% and 90% of teachers had fewer than 25 students in their class, whereas, in France and Slovenia, only 30.6% and 38.1% had fewer than 25 students in their class. In France, Ireland and Luxembourg, half of the teachers or more reported that between 25% and 49% of their students struggled at school, while in Slovenia and Italy most teachers reported that less than one quarter of their students had educational difficulties.

7551 students completed the questionnaire. While almost 90% of teachers’ classes were surveyed in Slovenia, only one quarter was surveyed in Italy and one fifth in France. The distribution of male and female students was equally balanced across samples. The distribution of students relative to teachers’ course subject was similar to that reported in the teacher sample, except in Ireland and Luxembourg where there were differences due to small sample sizes. In France and Ireland, most of the students surveyed were in grade 10. In Slovenia, students were enrolled in grades 9, 10 and 11, and a few in grade 12. It is worth noting that the Italian sample consists of older students – the majority of them being in grade 12. Finally, the Luxembourg sample comprises students in grades 9 and 11.

199 school leaders completed the questionnaire in France, Italy, Luxembourg and Slovenia. Irish results are not reviewed here due to the low number of responses. National samples differ in the distribution of school sizes. While 54.4% of schools in Italy and 80% of schools in Luxembourg have above 1000 students, only 35.9% of schools in France and 5.3% in Slovenia have such large numbers of students in their schools. On the other hand, 48% of Slovenian schools have fewer than 500 students. It is also worth noting that most of the schools selected in Italy, Luxembourg and Slovenia have participated in other projects related to digital technology in the last five years, while this is the case for only 35.9% of French schools. Furthermore, only 9.4% of French schools have participated in studies related specifically to artificial intelligence in the last five years, compared to 29.1% in Italy and 17.3% in Slovenia.

The 88 teachers interviewed in the five countries were predominantly female. 39 taught maths, 35 modern languages, and 14 another subject. 18 school leaders were interviewed in France, Italy and Slovenia. Due to the small number of volunteers for interviews in Ireland and Luxembourg, it was decided not to proceed with the interviewing process of school leaders in these two countries. Details of the sample characteristics are presented for each country in Table 8.

Co-funded by the
Erasmus+ Programme
of the European Union

Table 5: Characteristics of the teacher sample (surveys)

		France		Ireland[5]		Italy		Luxembourg		Slovenia		All countries	
		n	%	n	%	n	%	n	%	n	%	n	%
Number of respondents		180		14		275		10		257		736	
Gender	Female	98	54.4	13	93	193	70.2	3	30	203	79	510	69.3
	Male	78	43.3	1	7	81	29.4	7	70	51	19.8	218	29.6
	Prefers not to say/Other	4	2.2	0	0	1	0.4	0	0	3	1.2	8	1.1
Teaching experience	Less than 10 years	35	19.4	6	43	72	26.2	2	20	67	26.1	182	24.7
	More than 10 but less than 20 years	51	28.3	6	43	66	24	7	70	79	30.7	209	28.4
	More than 20 years	94	52.2	2	14	137	49.8	1	10	111	43.2	345	46.9
Subject	Maths	99	55	7	50	87	31.6	3	30	114	44.36	274	37.2
	Modern languages	73	40.6	6	43	98	35.6	5	50	92	35.8	310	42.1
	Computer science	0	0	0	0	37	13.4	0	0	10	3.89	33	4.5
	Experimental sciences	0	0	0	0	30	11	0	0	0	0	51	6.9
	Other	8	4.4	1	7	23	8.4	2	20	41	15.95	68	9.2
School type	Academic	111	61.7	8	57.1	155	56.4	5	50	93	36.2	372	50.5
	Vocational	20	11.1	4	28.6	101	36.7	5	50	112	43.6	242	32.9
	Other	17	9.4	2	14.3	19	6.9	0	0	52	20.2	90	12.2

[5] For the international report, only the responses from teachers who completed both the baseline and endline questionnaires were kept. However, to present a more comprehensive picture of Irish teachers' knowledge, perceptions and use of AI at the beginning of the experiment, the Irish evaluation report encompasses all teachers who completed the baseline questionnaire. Consequently, there may be disparities in findings between the international and Irish national reports owing to variations in sample composition.

Co-funded by the
Erasmus+ Programme
of the European Union

	Lower-secondary	32	17.8	0	0	0	0	0	0	0	0	32	4.3
Number of students in the class	0-24	55	30.6	9	64	217	78.9	9	90	98	38.1	388	52.7
	25-30	69	38.3	5	36	58	21.1	1	10	138	53.7	271	36.8
	More than 30	56	31.1	0	0	0	0	0	0	21	8.2	77	10.5
Percentage of students with educational difficulties in the class	0-24%	47	26.1	6	43	143	52	3	30	210	81.7	409	55.6
	25-49%	95	52.8	7	50	107	38.9	6	60	34	13.2	249	33.8
	50-100%	38	21.1	1	7	24	8.7	1	10	13	5.1	77	10.5
	NA	0	0	0	0	1	0.4	0	0	0	0	1	0.1

Table 6: Characteristics of the student sample

		France		Ireland		Italy		Luxembourg		Slovenia	
		n	%	n	%	n	%	n	%	n	%
Number of respondents		1133		92		1590		46		4690	
Gender	Female	570	50.3	43	47	714	44.9	20	43.5	2.427	51.7
	Male	487	43	42	56	831	52.3	22	47.8	1.918	40.9
	I prefer not to say/other	76	6.7	7	7	45	2.8	4	8.6	345	7.4
Subject of the teacher engaged in AI4T	Maths	573	50.6	29	32	574	36.1	24	52.2	2.123	45.3
	Modern languages	440	38.8	63	68	509	32	16	34.8	1.656	35.3
	Other	120	10.6	0	0	507	31.9	6	13	911	19.4
Grade	Grade 9	76	6.7	0	0	0	0	25	54.3	1.585	33.8
	Grade 10	900	79.4	92	100	310	19.5	1	2.2	1.679	35.8
	Grade 11	155	13.7	0	0	372	23.4	20	43.5	1.148	24.5
	Grade 12	0	0	0	0	908	57.1	0	0	257	5.5

Co-funded by the Erasmus+ Programme of the European Union

Table 7: Characteristics of the school leader sample (survey)

		France		Italy		Luxembourg		Slovenia	
		n	%	n	%	n	%	n	%
Number of respondents		64		55		5		75	
School size	<500 students	12	18.8	5	9.1	1	20	36	48.0
	500-999 students	29	45.3	20	36.4	0	0	35	46.7
	>1000 students	23	35.9	30	54.5	4	80	4	5.3
Percentage of schools who participated, in the last 5 years, in studies related to...	...Digital tools	23	35.9	39	70.9	4	80	60	80
	...AI	6	9.4	16	29.1	0	0	13	17.3

Table 8: Characteristics of the teachers and school leaders interviewed

		France	Ireland	Italy	Luxembourg	Slovenia
		n	*n*	*n*	*n*	*n*
Number of teachers		23	9	28	10	18
Teachers' gender	Female	9	8	22	3	15
	Male	14	1	6	7	3
Teachers' course subject	Maths	14	5	10	3	9
	Computer sciences / Experimental sciences	0	0	8	0	0
	Modern languages	9	3	10	5	8
	Other	0	1	0	2	1
Number of school leaders		5	0	7	0	6

Co-funded by the
Erasmus+ Programme
of the European Union

3.2 Data processing

Data cleaning

Due to the administration method that allowed for multiple responses coming from a single participant, the first step of the data cleaning process consisted in removing duplicates, which were identifiable thanks to the evaluation numbers entered by participants. If a single participant answered several times, the most complete answer was kept, and if several answers had the same level of completion, the first one was kept. Incomplete answers were kept if the participant had completed at least the first module of questions. For the teacher sample, responses were kept only when the teachers had answered both questionnaires.

The correspondence between participants' evaluation numbers, which were specific to every country, and the country entered by participants was checked. A few participants indicated a country that was not consistent with their evaluation number. In this case, their entry for the country was modified by the evaluator.

Psychometrics properties of the scales

Before calculating the scales scores, the scales psychometric properties were tested. The Cronbach alpha was calculated on all scales as a measure of internal consistency. For each item, the evaluation team calculated the item-total correlation and the *alpha if item is dropped*. Items were taken out of the scale when their correlation with the total was significantly lower than the other items and when their removal improved the alpha. A factor analysis was then conducted for each scale. We used Cattell's scree test to identify the number of factors. Additional items were taken out when we identified cross-loadings on several factors. A summary of the psychometric properties of the scales can be found in the appendix A.

To calculate the scores, the Likert scales were converted into numbers. The scores on each item were added together and then divided by the number of items. Standardisation was operated at the country level based on the mean and standard deviation of the control group in the baseline.

Balancing checks & attrition

Before conducting the impact analysis, the evaluation team checked that the randomisation had produced two comparable groups of teachers in each country. To this aim, we performed a student t-test on teacher characteristics and on the main outcomes measured at the beginning of the experiment. Observing significant differences between the two groups is likely in small samples such as ours. They do not invalidate the randomisation process, but they reinforce the importance of taking into account control variables in the regression analyses. The results of these analyses can be found in appendix B.

Compliance

At the end of the experiment, teachers were asked whether they had received access to the AI4T intervention resources. The results show that randomisation was successful in the countries where an impact evaluation was implemented. Only 7.4% of teachers from the control group in Slovenia, 5.1% in Italy, and 1.7% in France reported that they had had access to the AI4T professional learning pathway. On the other hand, 3.4% of teachers from the intervention group in France, 0.8% in Slovenia and 0.7% in Italy reported that they had not received access to the AI4T professional learning pathway.

Further details are provided on their actual engagement with the professional learning pathway in the section Completion and Engagement.

Qualitative data processing

Qualitative data was gathered via interviews of teachers and school leaders and via open-ended questions in the questionnaires:

1. The interviews with teachers and school leaders took place on Microsoft Teams or Zoom, after obtaining their consent to be interviewed and recorded.

2. Machine transcriptions were made (Zoom, NVivo) whenever possible, and each transcript was then amended to fix any errors. The recorded interviews will be deleted at the end of the project.
3. The research team created two analytical grids, one for the teacher interviews and another for the school leader interviews. This allowed for cross-country comparisons of the interview data. The grids were informed by the different aims of the research process, the theoretical framework and the interview protocol. They were tested on a small sample of interviews and amended when necessary.
4. Using NVivo or Taguette software, each transcript was analysed, assigning labels (or "codes") to each chunk of data, based on the analytical grid. *In Ireland, the team first coded each chunk of data without following any specific grid, then identified recurring themes informed by the shared analytical grid.*
5. The information collected was used to illustrate, confirm, explain, as well as shed some new light on the data collected via the questionnaires.

Common analytical grids were also devised at the international level to analyse the open questions.

4. Teachers' results

4.1 Teachers' reactions to the professional learning pathway

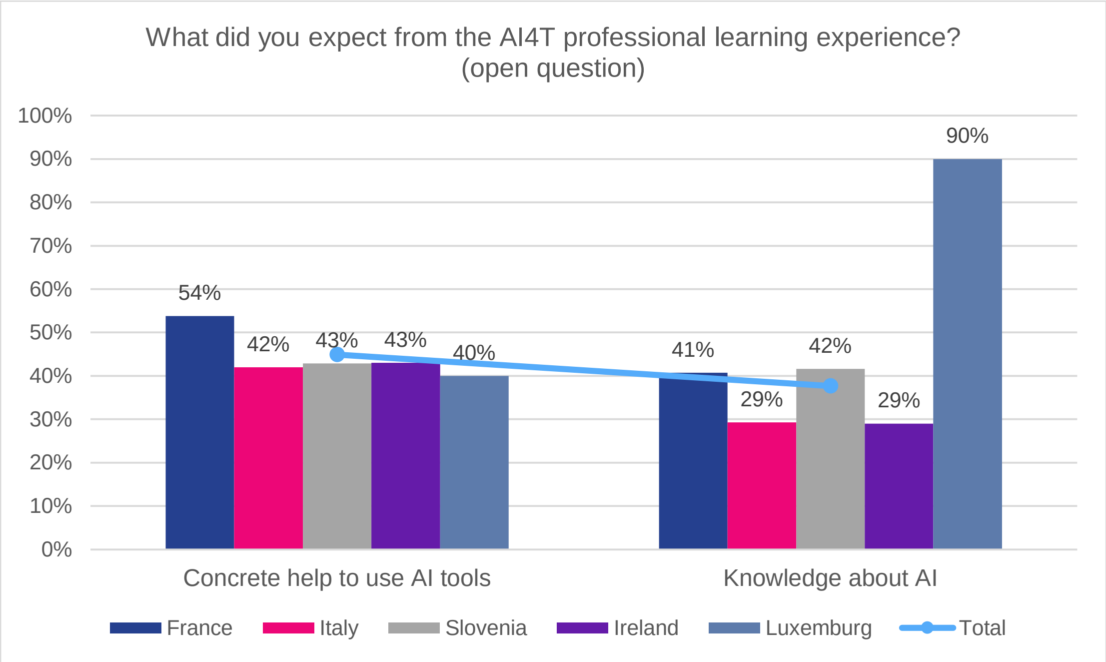
Expectations

At the end of the experiment, teachers were asked about their **expectations** via an open-ended question. Two main expectations were apparent in the five countries:

1. Teachers wanted to receive assistance in how to use AI tools. 42% of Italian teachers, 53.8% of French teachers and 63% of Slovenian teachers expressed this wish. 3 Irish teachers (43%) and 4 Luxembourgish teachers (40%) made similar statements. In Slovenia and France, many interviewees insisted on the need for **concrete** assistance and **real examples of use** in classroom situations.
2. Teachers expected to learn more about AI. 29.30% of Italian teachers, 40.7% of French teachers and 41.6% of Slovenian teachers expressed expectations related to learning about AI. It was also the case for 2 Irish (29%) and 9 Luxembourgish teachers (90%).

When asked if their expectations had been met, Italian teachers reported a high level of satisfaction with 43.4% of participants choosing the answer "completely" and 44.1% "for the most part". Similarly, in Luxembourg and Ireland almost all teachers (10 out of 10, and 6 out of 7 respectively) answered "completely" or "for the most part". In Slovenia, most teachers stated that their expectations had been met "for the most part" (55.8%). However, in France, teachers were more divided with 36.3% of teachers answering that their expectations had been met "for the most part" and 48.4% answering only "a little".

Figure 3: Teachers' expectations from the AI4T professional learning experience



Completion and engagement

Levels of completion of the MOOC, Webinars and face-to-face sessions were high in all five countries. Between 86.4% (Slovenia) and 100% (Ireland) completed the MOOC "partially" or "fully"; and between 89.2% (Italy) and 100% (Luxembourg) completed the webinars "partially" or "fully". 85.7% in France and 100% in Ireland and Luxembourg completed the face-to-face sessions "partially" or "fully". Completion of the textbook, however, differed from country to country. In Slovenia, there was a high level of completion, with 81.8% of teachers stating that they had "partially" or "fully" completed the textbook. The proportion was lower in Italy, but still represented 61.8% of teachers. In France, this particular resource was used by less than half of the sample (45.1%). Finally, 7 out 10 (70%)

Luxembourgish teachers and 4 out of 7 Irish teachers (57%) stated that they had used the textbook. The lower proportion in France may be attributed to the non-mandatory use of the textbook within the French professional learning pathway, leading to less communication encouraging its use. While the textbook was presented in February in Slovenia, it was presented a month later in France (only one week before administering the last questionnaire).

On average, participants reported medium to high levels of behavioural and cognitive engagement, with scores ranging from 4.72 to 5.17 on a scale of 1 to 7, respectively. Social engagement was relatively low, with a score of 3.66, while emotional engagement was high, with a score of 5.62. These results suggest that most participants actively engaged with the materials (taking notes and rereading their notes), made the effort to understand course content and enjoyed the learning experience. However, most of them did not participate regularly in course discussions or shared learning materials.

Social engagement is particularly low in France, with an average score of 3, compared to a score of 3.48 in Slovenia, 4.07 in Italy and 4.1 in Luxembourg. Finally, Ireland is the country with the highest level of social engagement at 5.81. These differences may be explained by the number of learners. The small sample size in Ireland and Luxembourg may have facilitated participation in course discussion. In Italy, webinars were also run in small groups while in France and Slovenia, webinars were run with all participants at once. In Italy, Luxembourg and Slovenia, participants were also requested to prepare activities and share them with others[6], which was not the case in France. In the interviews and open questions, French participants praised the interactions that took place during the face-to-face sessions (organised in small groups). However, this may not be reflected in the score as they were not *regular* interactions.

In all five countries, participants did not, on the whole, report difficulties in engaging with the professional learning pathway. No particular issues were reported in terms of equipment, room availability, online bugs, or support from the school administration.

Satisfaction

Satisfaction with all parts of the professional learning pathway was generally high across countries. Slovenian teachers showed consistent satisfaction levels towards the various components of the professional learning pathway. For example, 77.8% expressed satisfaction with the textbook and 74.8% with the MOOC. Italian and French teachers' satisfaction rates varied depending on the specific element of the learning pathway. Italian teachers showed slightly less enthusiasm for the textbook (63.9%) but higher satisfaction rates for the webinars (77.9%) and the MOOC (84.6%). French teachers were less satisfied with the webinars (46.6%) but were more positive towards the textbook (63.4%), the MOOC (79.8%) and the face-to-face sessions (80.8%). Teachers from Luxembourg and Ireland predominantly reported high levels of satisfaction, often reaching 100%, with the various components of the AI4T pathway.

When assessing the **usefulness** of the AI4T professional learning pathway for their work, Italian teachers gave an average score of 5.10 out of a 7-point Likert scale. Slovenian and French teachers provided lower scores of 4.75 and 4.01, respectively. The score of 4.01 from French teachers implies a neutral stance, neither agreeing nor disagreeing, in average, on the usefulness of the pathway in their work. Teachers from Ireland and Luxembourg expressed noticeable satisfaction, with mean scores of 5.76 and 5.7, suggesting a tendency towards agreement or general agreement regarding the usefulness of the pathway in their work.

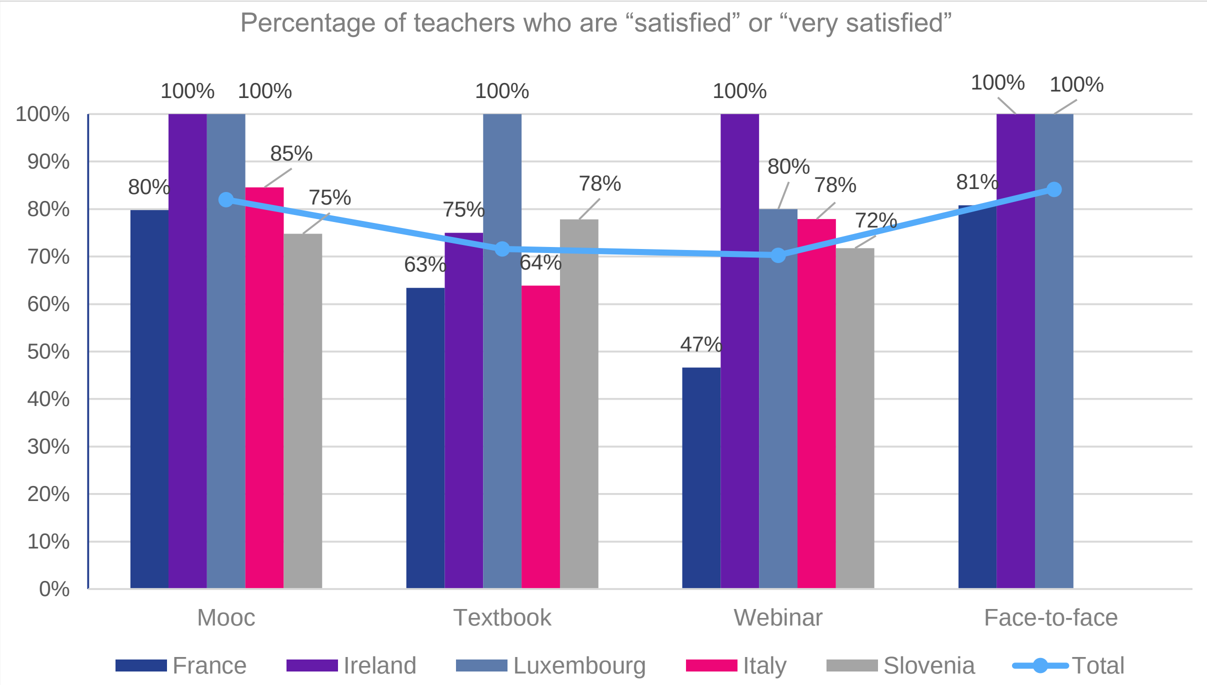
French teachers highlighted that the pathway did not have enough **practical value for their work** (only 30.8% stated being satisfied on this point), even though 51.6% were satisfied that the pathway had helped them **improve their professional skills**. On these two points, Slovenian teachers were more positive (53.2% / 61.7%), as were Italian teachers (65.8% / 76.3%). However, Slovenian teachers were

[6] This type of interaction was measured in the questionnaire and contributed to the score on social engagement.

less satisfied with the appropriateness of the pathway resources for their course **subject** (37.7%), as compared to French (57.1%) and Italian teachers (69.1%). French teachers were the most critical of how the way the content was taught enabled them to take an **active role** in the pathway (38.5%), and of the **possibilities offered to share professional experience with other participants** (47.3%), while Slovenian and Italian responses varied between 55.9% and 72.1% on these points.

Luxembourgish and Irish teachers were quite positive when assessing the practicality of the pathway. Almost all Luxembourgish and Irish teachers indicated that it helped them improve professional skills and was relevant to the subjects they taught. Almost all of them also appreciated the opportunities for interaction with fellow participants and for active involvement in the pathway.

Figure 4: Satisfaction with the professional learning pathway



In the interviews, many teachers expressed their **satisfaction with the blended format**, whether it provided a combination of online and face-to-face sessions (France, Ireland and Luxembourg), or large group online presentations and small-group seminars (Italy). Teachers highlighted the benefits of online resources and sessions (e.g. self-pace learning, no travel constraint), as well as their limitations (e.g. dissatisfaction with spending more time on screens), and they tended to highly value the interactions during the face-to-face sessions.

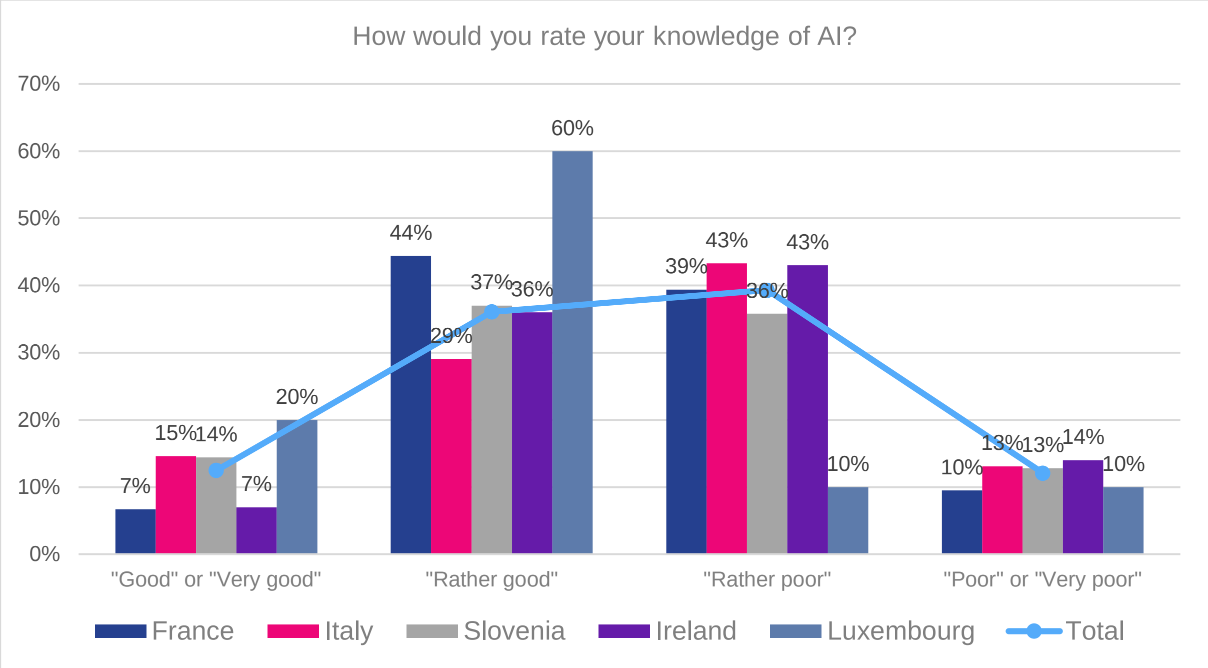
To conclude, it is worth noting that most participants across all countries expressed satisfaction with the **responsiveness of the training teams in relation** to participants' questions, with figures ranging from 80.2% in France to 100% in Luxembourg and Ireland.

4.2 Teachers' learning

Pre-experiment knowledge of AI

At the beginning of the experiment, teachers were asked to self-assess their knowledge of AI. Most teachers (75.4%) reported a **medium level of knowledge** choosing the answer options "rather poor" or "rather good". In Luxembourg, teachers estimated their level of knowledge of AI higher than in other countries: 8 teachers out of 10 chose the answer options "rather good" or "good".

Figure 5: Teachers' self-assessed knowledge of AI

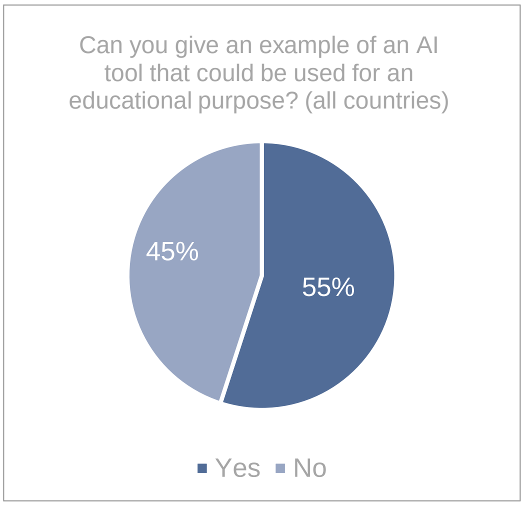


To assess their understanding of the concept, teachers were asked to provide a definition of AI. A response given by many participants was that AI is a software, programme or algorithm – 41.7% in France, 41.6% in Slovenia, 7 teachers in Ireland (50%) and 5 teachers in Luxembourg (50%). In Italy, only 13.1% identified AI as a software. This is due to participants favouring the term “machine” (“macchina”) to describe AI. However, almost half of the Italian teachers (48.7%) associated AI with the notion of intelligence, reasoning or the imitation of intelligence, compared to an average 30.2% across the five other countries. A minority of participants mentioned technical aspects such as the fact that AI collects data (7.2% across countries), AI processes data (19.4%), AI takes decisions (12.9%), or AI learns (23%). Some participants (15.6%) also mentioned elements associated with the goals of AI, such as AI “facilitates specific tasks” or “helps personalise teaching”. The proportion on this specific response was higher in France (32.8%) and Ireland (43%). Most teachers did not provide detailed definitions, with 60.3% of teachers overall mentioning only one or two of the response types listed above. Finally, 24% of teachers did not mention any of these responses.

Scores on familiarity with AI technologies, e.g. “machine learning”, “neural network”, “supervised learning”, were low in all countries – between 2.07 (France) and 2.32 (Luxembourg). This score is calculated on a scale of 1 to 5 and corresponds to the “rather unfamiliar” response option.

When presented with types of AI tools (e.g. machine translators, automated essay grading software) and asked if they indeed contained AI, teachers generally answered correctly. However, almost half of the teachers in Slovenia (41.3%), France (45.6%) and Italy (49.1%) reported at the beginning of the questionnaire that they could not give the name of an AI tool that could be used for an educational purpose. It was also the case for 5 Irish teachers (36%) and 3 Luxembourgish teachers (30%).

Figure 6: Teachers' ability to mention an AI tool for education



Impact

An impact analysis was conducted on the three countries with large enough sample sizes, namely France, Italy and Slovenia. Intervention impact was tested on five outcomes related to AI knowledge – **self-assessed knowledge of AI, knowledge of how AI works, familiarity with AI technologies, identification of AI in tools mainly based on AI, and identification of AI in tools not mainly based on AI.** In table 9, each line corresponds to one regression with the outcome as the dependent variable. The table presents only, for each of the five models, the regression coefficients of the variable "having participated in the intervention", which are reported in the column *Intervention impact*. However, the models also included the following control variables: randomisation (0: control; 1: intervention group), time (0: baseline; 1: endline), gender (0: female; 1: male), years of teaching experience, course subject (language), course subject (mathematics), school type (other type of school), school type (vocational), and self-efficacy for integrating technology in the classroom. For a detailed presentation of the models' results, see the national evaluation reports on the AI4T website.

The effect of the intervention is strong and significant across all countries for three variables, namely "self-assessed knowledge of AI", "familiarity with AI technologies", "identification of AI in tools mainly based on AI". For all three variables, intervention impact is more pronounced in France. The intervention yields significant effects on "knowledge of how AI works" in Slovenia and Italy. While not significant in France, the effect size closely aligns with those observed in other countries, suggesting that the lack of significance may be attributed only to the smaller sample size in France. Finally, intervention impact is significant only in Slovenia for the variable "identification of AI in tools that are not primarily based on AI." Interpreting this variable proved challenging due to the rapid evolution of technology, which made it difficult to discern right and wrong answers regarding the inclusion of AI in such tools.

Table 9: Impact of the intervention on teachers' knowledge of AI

Outcomes	*Intervention impact*	*Standard error*	*p-value*
France			
Self-assessed knowledge of AI	0.602***	0.180	<0.01
Knowledge of how AI works	0.323	0.213	n.s
Familiarity with AI technologies	1.241***	0.201	<0.01
Identification of AI in tools that are mainly based AI	0.873***	0.228	<0.01
Identification of AI in tools that are not mainly based on AI	0.320	0.199	n.s
Italy			
Self-assessed knowledge of AI	0.540***	0.139	<0.01
Knowledge of how AI works	0.380**	0.155	<0.05
Familiarity with AI technologies	0.918***	0.155	<0.01
Identification of AI in tools that are mainly based on AI	0.791***	0.164	<0.01
Identification of AI in tools that are not mainly based on AI	0.269	0.169	n.s
Slovenia			
Self-assessed knowledge of AI	0.471***	0.139	<0.01
Knowledge of how AI works	0.445**	0.197	<0.05
Familiarity with AI technologies	0.612***	0.160	<0.01
Identification of AI in tools that are mainly based on AI	0.700***	0.176	<0.01
Identification of AI in tools that are not mainly based on AI	0.597***	0.177	<0.01

Reading note: In France, teachers who participated in the intervention scored higher than teachers with similar characteristics who did not participate in the intervention, by 0,602 (60% standard deviation) in average on "self-assessed knowledge of AI".

In Luxembourg and Ireland, the smaller sample size did not allow for an impact analysis. However, similar trends were observed on the knowledge outcomes. In Ireland, 8 teachers out of 9 reported that their level of knowledge of AI was good or rather good after engaging with the professional learning pathway, compared to 4 teachers prior to engaging with the pathway. In Luxembourg, teachers had a higher level of knowledge to begin with but that level increased further after engaging with the professional learning pathway – 7 out of 10 teachers indicated that their knowledge was “good” or “very good,” compared to 2 teachers at the beginning of the experiment. In Ireland, teachers were also better able to recognise the presence of AI in tools such as machine translators. All Irish teachers recognised machine translators as AI by the end of the experiment while it was only the case for 4 out of 9 teachers at the beginning of the experiment.

In all five countries, the interviews also confirmed that participants generally considered that they had learned more about AI thanks to the professional learning pathway. Teachers from Luxembourg, Ireland and France highlighted that they now felt more confident recognising AI and that they understood that AI is present in software and digital tools to a greater extent than they initially thought. Teachers from both France and Ireland also emphasised that they were more confident talking to their students about AI.

4.3 Teachers’ perceptions of AI

Pre-experiment perceptions of AI

In all countries, participants generally started with a positive attitude towards AI. Teachers reported **a high level of AI enjoyment** with an average score among the five countries of **5.72** and a **low level of AI anxiety**: **2.83**. Both scores are calculated on a scale of 1 to 7.

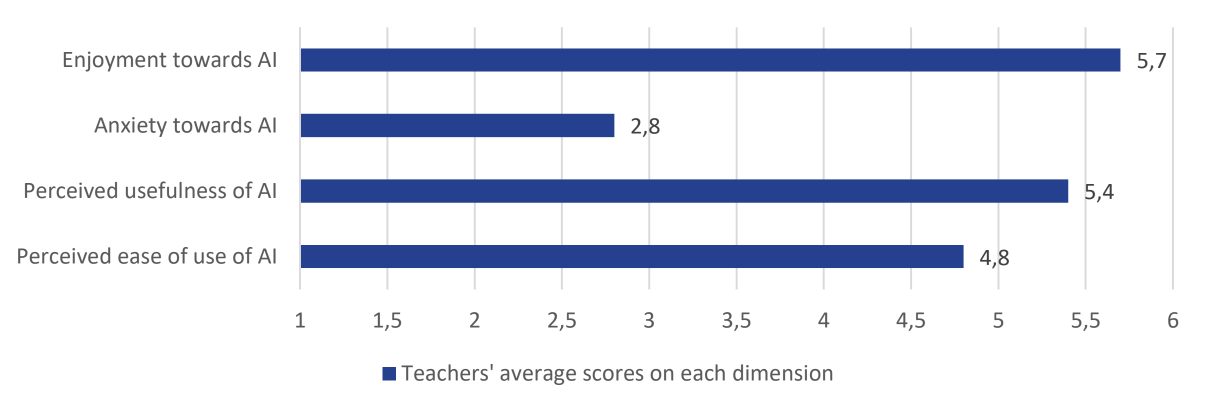
When asked to indicate emotions that they associated with AI, 48.2% of teachers mentioned emotions such as “curiosity”, “enthusiasm”, “interest”, etc., which were grouped in the “drawn to AI” category. In contrast, 25.7% mentioned emotions associated with apprehension towards AI, such as “fear” or “worry,” and 18.7% reported emotions associated with pleasure such as “ease” or “fun”. Although there is a general tendency in favour of positive emotions, there is a noteworthy difference between countries. While 65.1% of the Italian sample reported emotions categorised as “drawn to AI”, only 26.7% of the Slovenian sample reported those same emotions. The proportion is also high in Ireland and Luxembourg with 10 out of 14 teachers (71.4%) and 7 out of 10 teachers (70%) respectively, who mentioned emotions in this category. The proportion in France for this category is 48.9%, which is close to the average. The proportion of Italian teachers (28.7%) who mentioned emotions of pleasure is also higher than the average.

The perceived usefulness of AI was high among teachers from all countries, with an average score of **5.4** on a scale of 1 to 7. A high proportion of teachers perceived AI to be particularly useful for administrative work (84.2%), creating lessons, exercises or tests (87.4%), grading (85.1%), and for monitoring students’ learning progress and behaviour (86.7%).

When asked about potential consequences of using AI, teachers tended to agree more with positive statements such as “teaching quality will increase” (64%) or “teaching will be personalised to each students’ needs” (66.6%). Only a minority of teachers agreed with negative statements such as “the teaching profession will be devalued” (13.3%) or “teachers will be progressively replaced with AI” (9.9%). Concerns with the highest agreement related to private companies’ potential influence on schooling (48.4%), surveillance in school (45%), and data security (46.3%). Teachers from the Italian sample tended to agree with negative statements to a lesser extent than other countries, as only 29.8% agreed with the statement on surveillance, and 32.7% with the one on data security.

Finally, there was a **medium to high level of perceived ease of use of AI among participants.** On a scale of 1 to 7, scores ranged from 4.63 in Slovenia to 5.48 in Luxembourg.

Figure 7: Teachers' pre-experiment perceptions of AI



Impact

The impact analysis was conducted on the samples provided by France, Slovenia and Italy on four variables – perceived ease of use of AI, anxiety associated with the use of AI and learning about AI, enjoyment associated with the use of AI and learning about AI, and perceived usefulness of AI for education. We observed **no significant effect of the intervention on the four variables measured in relation to teachers' perceptions of AI, except for perceived ease of use of AI which increased significantly as a result of participating in the professional learning pathway in Italy.**

Table 10: Impact of the intervention on teachers' perceptions of AI

Outcomes	*Intervention impact*	*Standard error*	*p-value*
France			
Perceived ease of use of AI	-0.246	0.215	n.s
Anxiety associated with use of AI and learning about AI	-0.187	0.209	n.s
Enjoyment associated with use of AI and learning about AI	-0.090	0.222	n.s
Perceived usefulness of AI for education	0.173	0.215	n.s
Italy			
Perceived ease of use of AI	0.440***	0.147	<0.01
Anxiety associated with use of AI and learning about AI	-0.102	0.151	n.s
Enjoyment associated with use of AI and learning about AI	0.183	0.170	n.s
Perceived usefulness of AI for education	0.156	0.163	n.s
Slovenia			
Perceived ease of use of AI	-0.005	0.150	n.s
Anxiety associated with use of AI and learning about AI	-0.094	0.172	n.s
Enjoyment associated with use of AI and learning about AI	0.168	0.165	n.s
Perceived usefulness of AI for education	0.305	0.189	n.s

Reading note: In Italy, teachers who participated in the intervention scored higher than teachers with similar characteristics who did not participate in it, by 0,440 (44% standard deviation) in average on "perceived ease of use of AI".

In Ireland and Luxembourg, teachers' perceptions also seemed constant before and after engaging with the professional learning pathway, without any noteworthy variations on any of the four variables.

The absence of a significant effect of the intervention on teachers' perceptions in most countries needs to be considered alongside their already positive attitude towards artificial intelligence. In addition, the

AI4T intervention coincided with the launch of ChatGPT and subsequently widespread public interest in AI. The analysis showed a significant effect of the control variable "time", independent of the effect of the intervention, on teachers' perceptions. Attributed to time only, there was a significant drop in the perceived usefulness of AI in France by a 36% standard deviation and in Slovenia by a 44% standard deviation; a significant decrease in AI enjoyment in France by a 35% standard deviation and in Italy by a 31% standard deviation; and a significant increase in AI anxiety in France by a 31% standard deviation.

In line with these results, teachers seemed to have a slightly less positive view of the consequences of AI by the end of the experiment compared to the beginning. In the endline questionnaire, 50.5% of teachers agreed that thanks to AI "teaching quality would increase", compared to 64% in the baseline questionnaire, and 57.5% agreed that thanks to AI 'teaching would be personalised to each students' needs" compared to 66.6%. This drop is observed in all five countries. Teachers also agreed to a greater extent that because of AI "students' personal information would be more at risk of being breached and used at their expense", with 56.5% of teachers in agreement with the statement compared to 46.3% at the beginning. This trend was observed in all countries except Luxembourg where the proportion remained constant at 40%.

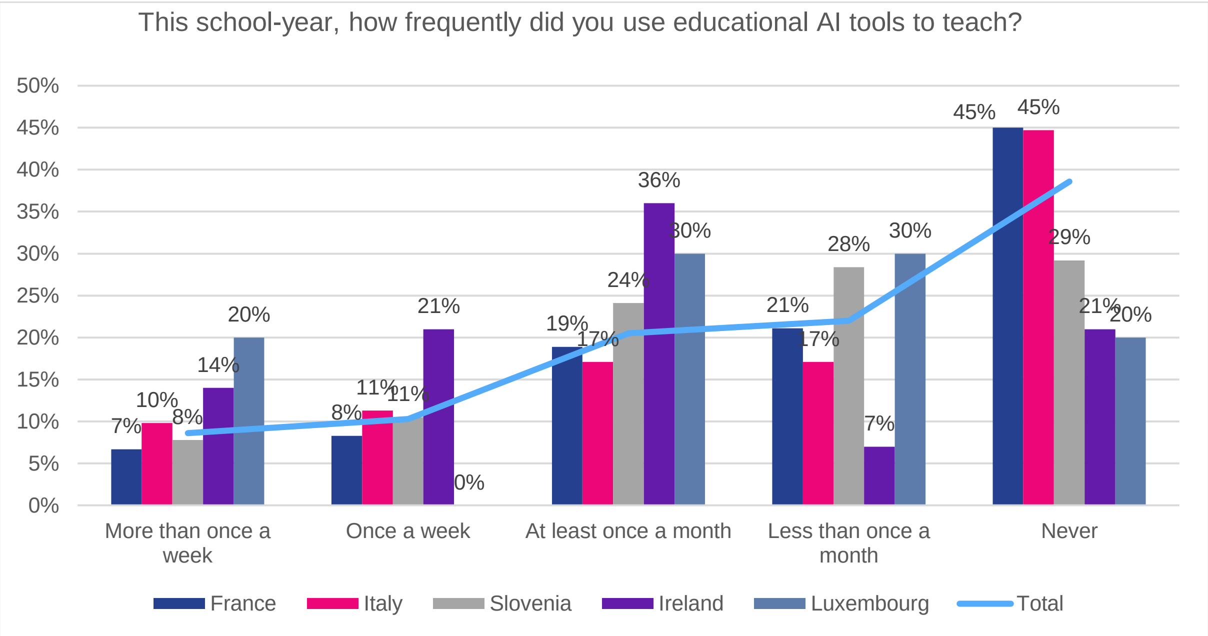
4.4 Teachers' intention to use AI and their actual use of AI

Pre-experiment intention to use AI and actual use of AI

At the beginning of the study, 38% of teachers stated that they hadn't used AI tools designed specifically for education that year. That proportion was lower in Slovenia representing 29.2% of teachers, in Ireland (3 out of 14 teachers) and in Luxembourg (2 out of 10 teachers), compared to France (45%) and Italy (44.7%). On the other hand, 18.9% of respondents stated that they used educational AI tools at least once a week, with only minor differences across countries, except in Ireland where 5 teachers out of 14 (35%) reported using AI tools on a weekly basis.

The higher use of educational AI tools in Slovenia and Ireland may partly be explained by the use of Photomath. Indeed, 33.3% of Slovenian maths teachers and 2 out of 7 Irish maths teachers (29%) mentioned using Photomath, as compared to 9.1% in France and 19.3% in Italy. In the Luxembourgish sample, the use of Duolingo was particularly high with 4 out of 6 language teachers (67%) claiming to use the tool.

Figure 8: Teachers' pre-experiment use of AI tools for education



However, many teachers also use AI tools that are not designed specifically for education. Almost all teachers use search engines and more than half of the language teachers (57.7%) use machine translators. Grammarly is also used by some language teachers in their job, namely 29.9% in Slovenia, 19.1% in Italy and 11% in France. Finally, intelligent personal assistants are used in Italy by 24.5% of language teachers, compared to 9.2% in Slovenia and 8.2% in France.

Teachers mainly use these tools for creating, presenting or sharing content with their students. 83% of teachers who claimed to use AI tools confirmed using them for this type of task. Almost half of the teachers (41%) also used AI tools for checking or marking exercises and tests. Less often, teachers use AI to monitor students' behaviour, work and learning progress (33%), or to diagnose their mistakes (12.3%).

Teachers were also asked about their ethical consciousness when using AI via statements such as "I am always aware of the abuse of AI technology" or "I have a good understanding of ethical issues arising from the use of AI tools". They reported a **medium to high level of ethical consciousness with a score of 5.1** on a scale of 1 to 7. Italian teachers reported a higher level of awareness than other countries with a score of 5.4.

Finally, almost all teachers (96.3%) stated that they would or probably would use AI tools during class sessions in the next five years. A similar proportion also stated that they would or probably would make students use AI tools in the next five years (96.2%).

Impact

The impact analysis did not show any significant effect on teachers' use of AI in France and in Slovenia. However, in Italy, the use of AI increased by a 42% standard deviation as a result of the intervention. We observed that in all three countries, there was no significant impact on frequent use of AI (teachers using AI more than once a week), indicating that, within the short time of the experiment, teachers did not integrate AI into their regular teaching practices, but instead that they merely tested AI tools. Finally, in Slovenia, there was a noteworthy 33% increase in standard deviation in teachers' intentions to use AI. Although the increase on this variable was higher in France, it was non-significant, likely due to the smaller sample size.

Table 11: Impact of the intervention on teachers' use of AI

Outcomes	*Intervention impact*	*Standard error*	*p-value*
France			
Use of AI	0.342	0.183	n.s
Frequent use of AI	0.186	0.204	n.s
Ethical consciousness when using AI	0.248	0.222	n.s
Intention to use AI	0.404	0.218	n.s
Italy			
Use of AI	0.417***	0.145	<0.01
Frequent use of AI	0.280	0.179	n.s
Ethical consciousness when using AI	0.314	0.177	n.s
Intention to use Ai	0.269	0.161	n.s
Slovenia			
Use of AI	0.279	0.144	n.s
Frequent use of AI	0.193	0.180	n.s
Ethical consciousness when using AI	0.307	0.189	n.s
Intention to use AI	0.333**	0.163	<0.05

Reading note: In Italy, teachers who participated in the intervention scored higher than teachers with similar characteristics who did not participate in the intervention, by 0,417 (42% standard deviation) in average on "use of AI".

Interviews in France and Slovenia conducted with teachers at the end of the experiment show greater nuance in teachers' intentions to use AI. In Slovenia, a specific analysis of teachers' interviews revealed different levels of readiness to integrate the tools in their teaching practice. 33.3% of the 19 teachers interviewed in Slovenia were moderately inclined to use AI tools. Indeed, they expressed interest but also a need for more information and training. 27.8% expressed a strong intention to use AI, and 22.2% were '*conditional adopters'* who stated that they would only use AI tools under specific conditions. 11.1%

of teachers were undecided as to whether to use AI tools, and 5.6% expressed scepticism about using the tools. The 'conditional adopter' profile was also encountered in the French evaluation results. Several French interviewees mentioned that they would only use AI if specific conditions were met, namely AI tools should be made freely accessible and GDPR-compliant, and they should meet their professional needs.

4.5 Impact variability

In this section, the evaluation team further explored whether the intervention had the same effect depending on teachers' actual engagement with the MOOC, teachers' course subjects and teachers' self-efficacy for integrating digital technology in the classroom. These analyses were conducted in France, Italy and Slovenia (countries with larger samples).

Engagement with the MOOC

The first analysis measures intervention impact on knowledge outcomes in two distinct groups of teachers, namely teachers showing above median levels of engagement with the MOOC, and teachers showing below median levels of engagement.

For all three countries, the impact is significant in both groups for most of the knowledge variables. In France and Slovenia, there is a coherent pattern whereby a greater engagement with the MOOC resulted in greater learning, found on almost all variables. In Italy, differences are smaller between the two groups and it has not been possible to identify any coherent pattern whereby one group has higher learning gains. This may be explained[7] in terms of the greater engagement with the MOOC displayed by all Italian teachers (see the report *Methodological Framework For Data Collection and Learning Analysis* (deliverable D1.3) for a more detailed analysis of teachers' engagement with the MOOC in each country).

Table 12: Impact of the intervention on teachers' knowledge of AI depending on their engagement with the MOOC

Outcomes	*Intervention impact*	*Standard error*	*p-value*
France			
Self-assessed knowledge of AI			
For teachers with higher MOOC engagement	0.658***	0.214	<0.01
For teachers with lower MOOC engagement	0.575***	0.213	<0.01
Knowledge of how AI works			
For teachers with higher MOOC engagement	0.591***	0.252	<0.01
For teachers with lower MOOC engagement	0.156	0.251	n.s
Familiarity with AI technologies			
For teachers with higher MOOC engagement	1.485***	0.227	<0.01
For teachers with lower MOOC engagement	1.206***	0.226	<0.01
Identification of AI in tools that are mainly based on AI			
For teachers with higher MOOC engagement	1.183***	0.261	<0.01
For teachers with lower MOOC engagement	0.913***	0.259	<0.01
Identification of AI in tools that are not mainly based on AI			

[7] To ensure the formation of two groups of equal size, teachers were divided based on the median value of the MOOC engagement indicator, rather than using a predetermined level of "high" and "low" engagement. As a result, it is possible for a group categorized with lower engagement to still exhibit a high level of engagement.

For teachers with higher MOOC engagement	0.346	0.235	n.s
For teachers with lower MOOC engagement	0.399	0.234	n.s
Italy			
Self-assessed knowledge of AI			
For teachers with higher MOOC engagement	0.450***	0.155	<0.01
For teachers with lower MOOC engagement	0.624***	0.155	<0.01
Knowledge of how AI works			
For teachers with higher MOOC engagement	0.394**	0.173	<0.05
For teachers with lower MOOC engagement	0.367**	0.173	<0.05
Familiarity with AI technologies			
For teachers with higher MOOC engagement	0.887***	0.174	<0.01
For teachers with lower MOOC engagement	0.948***	0.174	<0.01
Identification of AI in tools that are mainly based on AI			
For teachers with higher MOOC engagement	0.762***	0.184	<0.01
For teachers with lower MOOC engagement	0.818***	0.184	<0.01
Identification of AI in tools that are not mainly based on AI			
For teachers with higher MOOC engagement	0.308	0.189	n.s
For teachers with lower MOOC engagement	0.211	0.189	n.s
Slovenia			
Self-assessed knowledge of AI			
For teachers with higher MOOC engagement	0.449***	0.156	<0.01
For teachers with lower MOOC engagement	0.472***	0.156	<0.01
Knowledge of how AI works			
For teachers with higher MOOC engagement	0.647***	0.221	<0.01
For teachers with lower MOOC engagement	0.296	0.221	n.s
Familiarity with AI technologies			
For teachers with higher MOOC engagement	0.664***	0.180	<0.01
For teachers with lower MOOC engagement	0.575***	0.180	<0.01
Identification of AI in tools that are mainly based on AI			
For teachers with higher MOOC engagement	0.792***	0.198	<0.01
For teachers with lower MOOC engagement	0.607***	0.198	<0.01
Identification of AI in tools that are not mainly based on AI			
For teachers with higher MOOC engagement	0.615***	0.198	<0.01
For teachers with lower MOOC engagement	0.557***	0.198	<0.01

Reading note: In France, teachers who participated in the intervention with high MOOC engagement scored higher than teachers with similar characteristics who did not participate in the intervention, by 0,658 (66% standard deviation) in average on "self-assessed knowledge of AI"; while teachers who participated in the intervention with low MOOC engagement only scored higher by 0,575 (58% standard deviation) in average.

Teachers' course subject

The second analysis measures intervention impact on teachers' use of AI tools and intention to use AI tools for maths and language teachers separately. **It is worth noting that the impact of the intervention on the use of AI tools is significant for maths teachers but not for language teachers across the three countries.** The effect of the intervention on intention to use AI tools is also significant solely among maths teachers in Slovenia, while it is non-significant in other countries regardless of the teacher's course subject.

The variability observed across the two groups might be attributed to differences in the tools available in each respective subject. In France, the language teachers who were interviewed mentioned that they did not have access to new AI tools, unlike maths teachers who confirmed effective access to a subject-specific AI tool (Kwyk) for the duration of the AI4T project. Consequently, many language teachers interviewed in France emphasised the lack of AI tools available for trial and use with their students as a major hindrance. They explained that tools discussed during the professional learning pathway either required a license fee or were prohibited by their management due to GDPR non-compliance. Another potential explanation may rest with the level of alignment between the professional learning pathways and the needs of each subject. In Italy, some interviewees perceived that the professional learning pathway was better suited to maths teachers than to language teachers. In Slovenia, it was difficult to draw any conclusions as a few interviewees from both maths and language subjects expressed the belief that the professional learning pathway was better suited to the other subject.

Table 13: Impact of the intervention on teachers' use of AI depending on their course subject

Outcomes	*Intervention impact*	*Standard error*	*p-value*
France			
Use of AI			
For modern languages teachers	0.086	0.228	n.s
For maths teachers	0.452**	0.202	<0.05
Intention to use AI			
For modern languages teachers	0.249	0.273	n.s
For maths teachers	0.444	0.242	n.s
Italy			
Use of AI			
For modern languages teachers	0.271	0.189	n.s
For maths teachers	0.775***	0.194	<0.01
Intention to use AI			
For modern languages teachers	0.381	0.212	n.s
For maths teachers	0.322	0.217	n.s
Slovenia			
Use of AI			
For modern languages teachers	0.070	0.179	n.s
For maths teachers	0.377**	0.172	<0.05
Intention to use AI			
For modern languages teachers	0.373	0.203	n.s
For maths teachers	0.387**	0.195	<0.05

Reading note: In Slovenia, maths teachers who participated in the intervention scored higher than math teachers with similar characteristics who did not participate in the intervention, by 0,377 (38% standard deviation) in average on "use of AI"; while modern languages teachers who participated in the intervention did not score significantly higher on this indicator than other modern languages teachers with similar characteristics.

Self-efficacy for integrating technology

Finally, the impact of the intervention on knowledge and use of AI is presented separately for two groups of teachers: those with above median levels of self-efficacy for integrating digital technologies in the classroom, and those with below median levels of self-efficacy.

In France, the effect of the intervention remained significant for both groups of teachers on the three learning outcomes where there was an impact in the general sample, i.e. self-assessed knowledge of AI, familiarity with AI technologies and identification of AI in tools mainly based on AI. For those three outcomes, teachers' learning gains were higher in the group with lower levels of self-efficacy.

A similar trend showing higher learning gains for teachers with lower self-efficacy is also observable in the Slovenian sample. In this sample, the impact of the intervention was significant on teachers' self-assessed knowledge of AI and knowledge of how AI works, solely for teachers with lower levels of self-efficacy. The impact is significant for both groups on the other learning outcomes.

In Italy, the impact is significant for both groups on all variables except *knowledge of how AI works* for teachers with lower self-efficacy. Overall, the two groups only displayed minor differences in terms of intervention impact.

Table 14: Impact of the intervention on teachers' knowledge of AI depending on their self-efficacy for technology integration in the classroom

Outcomes	*Intervention impact*	*Standard error*	*p-value*
France			
Self-assessed knowledge of AI			
For teachers with higher self-efficacy for technology integration	0.500**	0.201	<0.05
For teachers with lower self-efficacy for technology integration	0.729***	0.212	<0.01
Knowledge of how AI works			
For teachers with higher self-efficacy for technology integration	0.270	0.238	n.s
For teachers with lower self-efficacy for technology integration	0.390	0.251	n.s
Familiarity with AI technologies			
For teachers with higher self-efficacy for technology integration	1.173***	0.224	<0.01
For teachers with lower self-efficacy for technology integration	1.324***	0.237	<0.01
Identification of AI in tools that are mainly based on AI			
For teachers with higher self-efficacy for technology integration	0.658**	0.253	<0.05
For teachers with lower self-efficacy for technology integration	1.142***	0.267	<0.01
Identification of AI in tools that are not mainly based on AI			
For teachers with higher self-efficacy for technology integration	0.426	0.223	n.s
For teachers with lower self-efficacy for technology integration	0.187	0.235	n.s
Italy			
Self-assessed knowledge of AI			
For teachers with higher self-efficacy for technology integration	0.516***	0.153	<0.01
For teachers with lower self-efficacy for technology integration	0.573***	0.163	<0.01
Knowledge of how AI works			
For teachers with higher self-efficacy for technology integration	0.440***	0.170	<0.01
For teachers with lower self-efficacy for technology integration	0.300	0.181	n.s
Familiarity with AI technologies			
For teachers with higher self-efficacy for technology integration	0.897***	0.171	<0.01
For teachers with lower self-efficacy for technology integration	0.945***	0.182	<0.01
Identification of AI in tools that are mainly based on AI			
For teachers with higher self-efficacy for technology integration	0.846***	0.181	<0.01
For teachers with lower self-efficacy for technology integration	0.718***	0.193	<0.01
Identification of AI in tools that are not mainly based on AI			
For teachers with higher self-efficacy for technology integration	0.285	0.186	n.s
For teachers with lower self-efficacy for technology integration	0.249	0.199	n.s
Slovenia			
Self-assessed knowledge of AI			
For teachers with higher self-efficacy for technology integration	0.292	0.158	n.s
For teachers with lower self-efficacy for technology integration	0.632***	0.154	<0.01
Knowledge of how AI works			
For teachers with higher self-efficacy for technology integration	0.113	0.223	n.s
For teachers with lower self-efficacy for technology integration	0.745***	0.218	<0.01
Familiarity with AI technologies			
For teachers with higher self-efficacy for technology integration	0.487***	0.183	<0.01
For teachers with lower self-efficacy for technology integration	0.725***	0.179	<0.01
Identification of AI in tools that are mainly based on AI			

For teachers with higher self-efficacy for technology integration	0.705***	0.201	<0.01
For teachers with lower self-efficacy for technology integration	0.695***	0.197	<0.01
Identification of AI in tools that are not mainly based on AI			
For teachers with higher self-efficacy for technology integration	0.643***	0.202	<0.01
For teachers with lower self-efficacy for technology integration	0.556***	0.198	<0.01

Reading note: In Slovenia, teachers with low self-efficacy for technology integration who participated in the intervention scored higher than teachers with similar characteristics who did not participate in the intervention, by 0,632 (63% standard deviation) in average on "self-assessed knowledge of AI". Teachers with high self-efficacy for technology integration who participated in the intervention did not score significantly higher on this indicator than other teachers with similar characteristics.

In France, no significant impact was observed on the use of AI tools and intentions to use AI tools variable, even when teachers were categorised according to their level of self-efficacy for integrating technology in the classroom. However, on the use of specific tools by maths teachers, such as Kwyk and Vittascience, there was an increase solely among teachers with higher levels of self-efficacy. In Slovenia, on the other hand, there was a significant impact on the variables *use* and *intention to use* solely for teachers with lower levels of self-efficacy. In Italy, the impact of the intervention on teacher's use of AI was significant for all teachers regardless of their level of self-efficacy, while the impact on intentions to use was non-significant for both groups.

Table 15: Impact of the intervention on teachers' use of AI depending on their self-efficacy for technology integration in the classroom

Outcomes	*Intervention impact*	*Standard error*	*p-value*
France			
Use of AI			
For teachers with higher self-efficacy for technology integration	0.391	0.205	n.s
For teachers with lower self-efficacy for technology integration	0.280	0.216	n.s
Use of Kwyk			
For teachers with higher self-efficacy for technology integration	0.263**	0.119	<0.05
For teachers with lower self-efficacy for technology integration	0.041	0.125	n.s
Use of Vittascience			
For teachers with higher self-efficacy for technology integration	0.199**	0.094	<0.05
For teachers with lower self-efficacy for technology integration	0.094	0.099	n.s
Intention to use AI			
For teachers with higher self-efficacy for technology integration	0.452	0.244	n.s
For teachers with lower self-efficacy for technology integration	0.345	0.257	n.s
Italy			
Use of AI			
For teachers with higher self-efficacy for technology integration	0.390**	0.159	<0.05
For teachers with lower self-efficacy for technology integration	0.453***	0.170	<0.01
Intention to use AI			
For teachers with higher self-efficacy for technology integration	0.272	0.177	n.s
For teachers with lower self-efficacy for technology integration	0.265	0.189	n.s
Slovenia			
Use of AI			
For teachers with higher self-efficacy for technology integration	0.235	0.165	n.s
For teachers with lower self-efficacy for technology integration	0.319**	0.161	<0.05
Intention to use AI			
For teachers with higher self-efficacy for technology integration	0.179	0.186	n.s
For teachers with lower self-efficacy for technology integration	0.473***	0.181	<0.01

Reading note: In France, teachers with high self-efficacy for technology integration who participated in the intervention scored higher than other teachers with similar characteristics who did not participate in the intervention, by 0,263 (26% standard deviation) in average on "use of Kwyk"; while teachers with low self-efficacy for technology integration who participated in the intervention did not score significantly higher on this indicator than other teachers with similar characteristics.

5. School leaders' results

The following section presents the results of the school leader evaluation conducted in France, Italy, Luxembourg and Slovenia at the end of the experiment. Due to the low number of school leaders who participated in the evaluation in Ireland, no results are presented for this country.

5.1 Schools' technical infrastructure

In all countries, school leaders reported **a good level of ICT equipment in schools** and, in France, the number of ICT devices available for student use was particularly high.

Across all countries, 81.9% of school leaders who participated in the evaluation reported that almost all **teachers in their schools were equipped with an ICT device** (desktop computer, tablet or laptop) that they could use in class. Italy scored the lowest at 66.7%, compared to 4 out of 5 (80%) in Luxembourg, 88% in France and 93.3% in Slovenia. In addition, 81.4% of school leaders (all countries) reported that the **internet connection** in classrooms was good overall. Regarding technical issues, almost all school leaders reported that teachers had access to **ICT support**. However, only 10.9% indicated that the support was available within the hour in France, as opposed to 64.8% in Italy and 42.7% in Slovenia.

68.8% of French schools and 4 schools out of 5 in Luxembourg (80%) provided **at least one ICT device (laptop, tablet or desktop computer) for every three students**, but this figure dropped to 27.3% in Italy and 14.6% in Slovenia. In most cases, available equipment could be used directly in the classroom. 66.3% of school leaders confirmed that existing ICT devices could be used directly in the classroom, while 30.2% expressed that it could be used in the classrooms only partly. On the other hand, only 32.2% of school leaders across all countries stated that the equipment could be used by students at home. The proportion rose to 48.4% in France, compared to 29.3% in Slovenia, 1 out of 5 (20%) in Luxembourg, and 18.5% in Italy. This shows that there were important differences in terms of student access to school-provided technology at home across participating schools.

5.2 Support for professional learning

Overall, school leaders who took part in the evaluation seemed to have encouraged teachers to take part in the AI4T project. A high percentage, 93.3% in Slovenia, 89.1% in Italy, and 73.4% in France, reported doing so. In schools belonging to the intervention group, most school leaders also reported having **provided teachers with information** about the professional learning pathway (97.6% in Italy, 83% in Slovenia, and 73.8% in France).

However, there were **important differences in the extent to which school leaders followed up on the professional learning pathway by engaging in discussions with teachers regarding their satisfaction with the pathway and by dedicating time to peer-to-peer dissemination of what teachers learned**. The great majority of school leaders in Slovenia (89.4%) and in Italy (82.9%) held discussions with teachers over their satisfaction with the professional learning pathway, compared to 50% of school leaders in France, and one out of three school leaders in Luxembourg. School leaders reported that time was devoted to peer-to-peer dissemination of learning gains in most schools in Slovenia and in almost half of the schools in Italy (61.7% and 48.8% respectively), but much less so in France (14.3%), and Luxembourg (one in three).

Most teachers did not receive **compensation** for their participation, although there were differences between countries. While 95.1% of Italian school leaders reported that teachers in their schools had not been paid for the hours they invested in the project, only 59.6% of Slovenian teachers, and one out of three school leaders in Luxembourg made the same statement. In France, 76.2% of school leaders reported that teachers had not been paid and 21.4% admitted to not knowing whether they had been paid. There were also differences between countries when it came to teachers having a **replacement teacher in the classroom** whilst they were participating in the professional learning pathway during teaching hours. 69% of teachers in France had not been replaced at all, in comparison to 2.1% of

Slovenian teachers. In Italy, 80.5% of school leaders reported that the professional learning pathway had not taken place during teaching hours. Concerning the refund of **expenses** related to the A4T experiment, school leaders in most countries claimed that teachers either had no expenses related to the AI4T professional learning pathway or that they had received a refund, except in France where almost half of the school leaders (45.2%) answered that they did not know whether teachers had received a refund.

5.3 AI leadership

School leaders' knowledge and use of AI

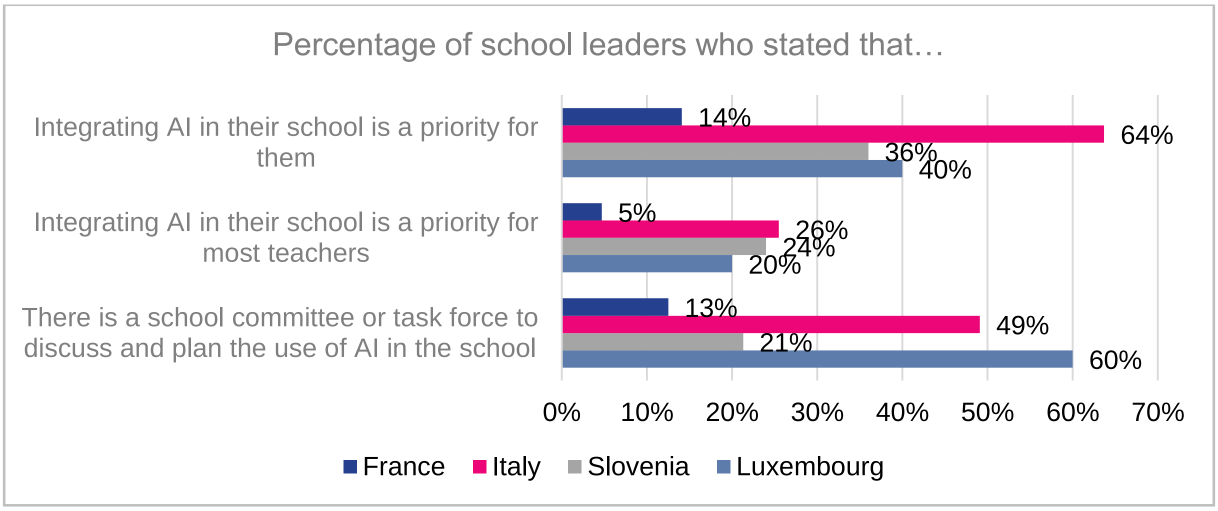
In Slovenia and France, school leaders reported lower levels of **knowledge of AI** than the teachers did. 56% of Slovenian school leaders and 64.1% of French school leaders self-assessed their level of AI knowledge as "rather poor" or lower. In Italy, by contrast, 65.5% of school leaders assessed their AI knowledge as "rather good" or higher. In France, most school leaders could not name an AI tool for teaching and learning (71.9%), as opposed to 34.7% in Slovenia and 45.5% in Italy. In the three countries, most school leaders stated that they do not use AI tools in their work (81.2% in France, 76% in Slovenia and 67.3% in Italy). At the time of the interviews, several school leaders mentioned that they had tested ChatGPT for their work.

School policy on AI integration

Overall, **a lack of AI leadership within the schools in Slovenia and in France was observed**. Only 14.1% of school leaders in France and 36% in Slovenia thought that integrating AI in their school was a priority. Similarly, only 12.5% of French schools and 21.3% of Slovenian schools had a committee or task force to discuss the use of AI. Most school leaders also believed that AI was not a priority for teachers. Only 14% of French school leaders and 24% of Slovenian school leaders thought that integrating AI was a priority for most teachers in their schools. Most school leaders (84.4% in France and 78.7% in Slovenia) did not spend any time planning, maintaining or administering the use of AI in their school. Furthermore, about half of the school leaders (59.4% in France and 46.7% in Slovenia) did not know whether teachers in their school had access to educational AI tools.

By contrast, in Italy, most school leaders thought that integrating AI in their school was a priority (63.7%) and about half of the schools (49%) had set up a committee or task force to discuss the use of AI. Indeed, 61.8% stated having spent three days or more planning, maintaining or administering the use of AI in their school. Most Italian school leaders (52.7%) also claimed that teachers in their schools had access to AI tools designed specifically for education.

Figure 9: School policy on AI integration



School leaders in the three countries held different views on the subject of **ethical issues** and AI integration in schools. Overall, Slovenian and French school leaders considered ethical issues to be important when deciding to integrate AI into the schools – with 50.7% and 53.1% stating that it played a major role or was the most decisive factor in AI integration. This figure dropped to 36.4% in Italy. Italian school leaders judged ethical considerations as playing an "average role" instead (45.5%). Nonetheless, most Italian (56.4%) and Slovenian (61.3%) school leaders claimed to have communicated with their staff on the subject of ethical guidelines regarding AI and data protection, as opposed to 34.4% in France. These figures were lower when it came to communicating with parents on those issues (about one quarter in Italy and Slovenia and 9.4% in France). Insights from interviews suggest that some French school leaders may not have disseminated any guidelines because they perceived AI integration to still be at an early stage.

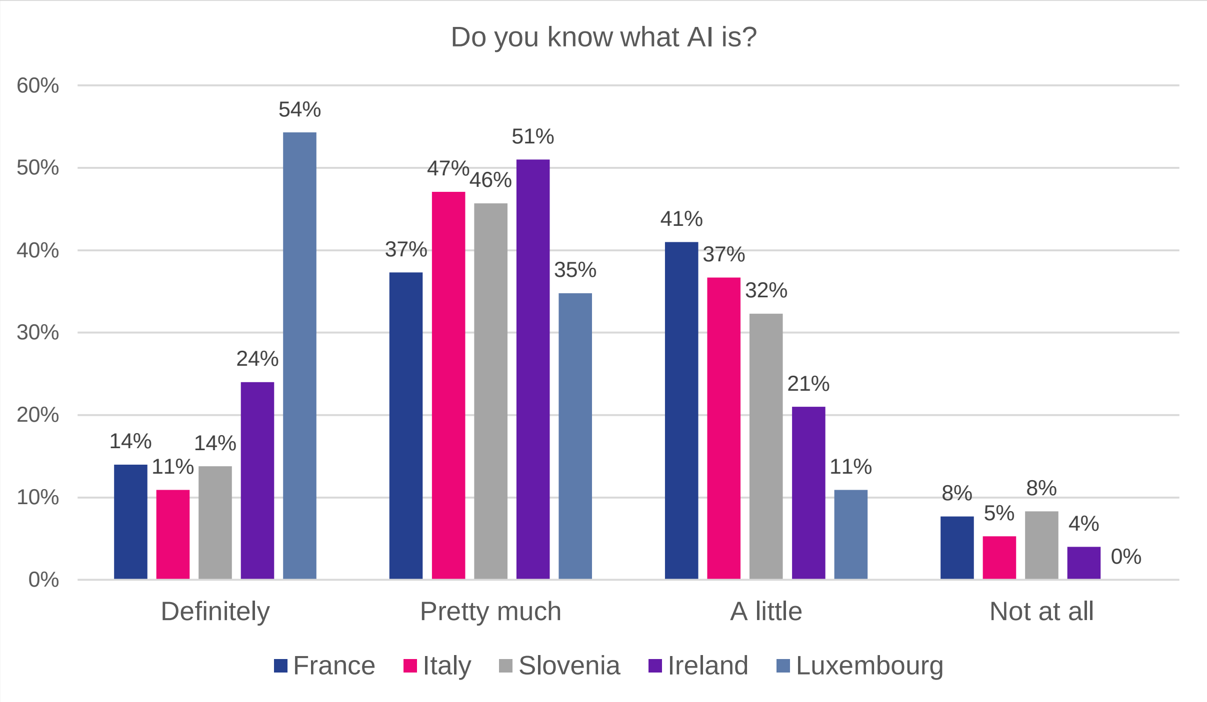
6. Students' results

The following section presents the results of the student evaluation conducted in the five countries, at the end of the experiment.

6.1 Students' knowledge of AI

First, students' understanding of artificial intelligence was assessed. When asked whether they knew what artificial intelligence was, very few students, from 0% (in Luxembourg) to 8.3% (in Slovenia), answered "not at all." Most students chose the answer option "pretty much" or "a little", except in Luxembourg where 25 students out of 46 (54.3%) answered "definitely".

Figure 10: Students' knowledge of AI



Overall, students' ability to recognise AI tools was good. The proportion of students who identified machine translators as AI ranged from 67.1% in Slovenia to 80.6% in France. Similarly, between 80.4% of students in Luxembourg and 88.7% in Italy recognised image-recognition systems as AI.

Students were also asked to define AI. The analysis of students' responses in Slovenia and Ireland showed a more nuanced picture. **Predominant themes in the Slovenian responses included: portraying AI as smart devices or software** (i.e. smartphones, apps, computer programs, algorithms, robots) – mentioned by 37.5% of students – or as **intelligent assistants** that can provide intelligent answers, make decisions, address goals or help with various tasks – mentioned by 24.5% of students. Similarly, in Ireland, 21.1% of students' responses included descriptions of AI as software or smart devices, and 27.8% described AI as digital assistance.

Just under a quarter (23.3%) of Irish students and 12.3% of Slovenian students also gave definitions of AI as **human-like intelligence**, referring to aspects of human cognition such as thinking, reasoning, learning, planning and problem-solving.

Less often, students referred to technical aspects of AI such as **a system that learns** (9.3% of responses in Slovenia and 11.1% in Ireland), or **a system that collects and processes data** (7.6% of responses in Slovenia and 11.1% in Ireland). The thematic analysis suggests that students had a practical, rather than a theory-based or technical, understanding of AI.

A more detailed analysis of the Slovenian data showed that **students rarely gave detailed definitions** containing several of the above-mentioned elements. In Slovenia, 30.3% of students' answers fell into one category, 20.9% into two categories, and only 6% spanned more than two categories. Finally, 42.8% of Slovenian students and 23.3% of Irish students gave vague or unrelated responses that could not be categorised.

7. Students' attitude towards AI

Overall, students exhibited a **tendency to either concur with positive statements and to remain neutral or express divided opinions in response to negative statements**. The average scores on the positive attitude scale (rated out of 5) ranged from 3.57 (France) to 3.99 (Italy), while the average scores on the negative attitude scales ranged from 2.91 (Italy) to 3.29 (Slovenia). Although differences between countries were small, Italy stood out as the country with both the highest score on the positive attitude scale and the lowest score on the negative attitude scale, suggesting a more positive attitude towards AI among Italian students.

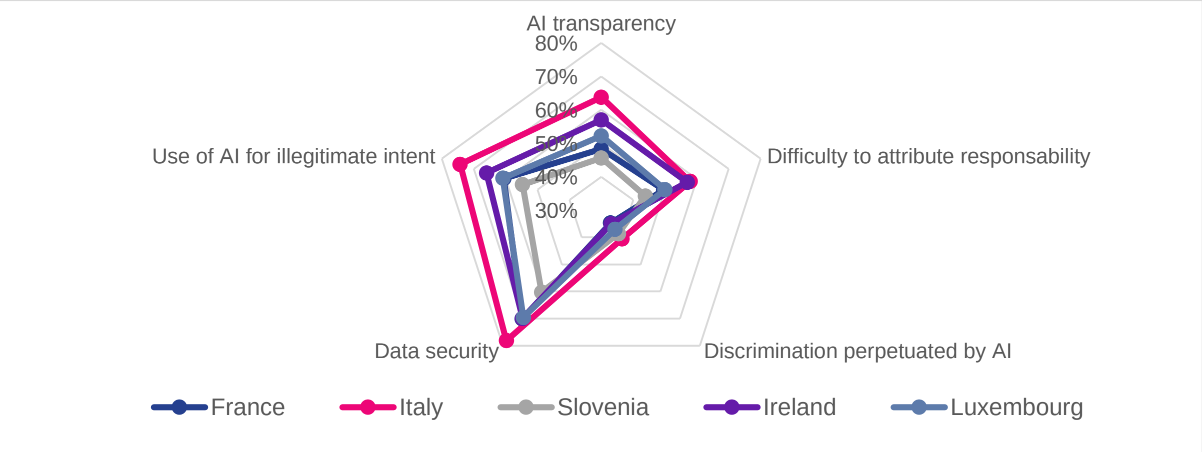
Between 70% (Slovenia) and 91.8% (Italy) of students stated that they were impressed with what AI can do. Most students also thought that AI would be useful for education, ranging from 61.5% in France to 80% in Ireland. Consequently, between 58.7% (Luxembourg) and 83.6% (Italy) of students were interested in discovering new AI tools for learning. However, nearly half of the students also agreed with negative statements such as "AI worries me" (agreement ranged from 30.6% in Italy to 52% in Ireland), or "I think the use of AI will lead to a greater risk of students' personal information being breached and used at their expense", from 41.3% in Luxembourg to 53.8% in Slovenia and France.

8. Students' ethical consciousness and worries regarding AI

Students were more aware of some ethical issues. From 60.3% (Slovenia) to 78.2% (Italy) of students reported being aware of the debate on **potential privacy violations due to data collection by AI tools,** and between 54.7% (Slovenia) and 74.3% (Italy) of students were aware of the debate on the **potential use of AI for illegitimate intents**. On the other hand, only around half of the students were aware of the debates on **AI transparency** or the **attribution of responsibility when AI makes decisions**. Finally, less than half of the students – between 34.7% (France) and 40.6% (Italy) – claimed to be aware of the debate on **potential discrimination perpetuated by AI tools**.

Italian students consistently demonstrated the highest awareness levels across all questions. This could be attributed to the older age composition of the Italian sample.

Figure 11: Students' awareness of ethical issues associated with AI



Students were also asked about their level of concern regarding each ethical issue. The results were consistent with the first question, with students reporting the highest level of concern over the issue of data privacy – between 63% (Ireland) and 70.7% (Italy) of students were "definitely" or "pretty much" concerned about this topic – and illegitimate use of AI – between 50% (Luxembourg) and 74.3% (Italy) of students were "definitely" or "pretty much" concerned about this issue.

9. Students' use of AI

At the end of the experiment, **most students reported that they had used AI tools with the teacher involved in the project**.

Less than one fifth of students in all countries said that they had never used generic AI tools with their teacher. The proportion ranged from 2% in Ireland to 16.3% in Slovenia. The proportion of students who had never used AI tools designed specifically for education (Photomath, Duolingo, Kwyk…) was higher in Slovenia (30.8%), in France (27.3%) and Italy (23.8%), than in Luxembourg (2.2%), and in Ireland (2%).

The proportion of students who claimed having used generic AI tools at least once a week varied greatly across samples: 33.8% in Slovenia, 43.5% in Luxembourg, 49.4% in France, 55.1% in Italy and 79% in Ireland. Similarly, the proportion of students who used educational AI tools at least once a week with their teacher differed across countries, from 23.7% in Slovenia to 64% in Ireland. These variations may be party explained by differences in sample size[8] and response rate[9].

In all countries, most students whose language teacher was involved in the project had used **machine translators,** ranging from 55% in Italy to 72.5% in France. **Duolingo** for schools was used by Irish students especially (95%), while it was used by only 18.9% of students in Italy, 10.3% in Slovenia, 9.5% in France and 6.2% in Luxembourg. **Smart assistants** (e.g Alexa, Siri, Cortana) were also used by students whose teacher was involved in the project. The proportion ranged from 15.5% in France to 25% in Luxembourg.

For students of maths teachers, **Photomath** was mostly used in Ireland (62%), Slovenia (56.7%) and Italy (34.7%), while **Kwyk** was mostly used in France (62.7%). In Luxembourg, 12.5% of maths teachers' students used **Checkmath**.

[8] It is important to consider that, although 92 students in Ireland and 46 students in Luxembourg participated in the survey, these samples were drawn from a very small number of classes.
[9] While almost 90% of teachers' classes were surveyed in Slovenia, only one quarter was surveyed in Italy and one fifth in France.

7. Takeaways from teachers and school leaders

This section outlines the takeaways from the evaluation which could inform educational policies on AI. It combines teachers' and school leaders' suggestions – collected through interviews – together with the quantitative findings.

7.1 On professional learning about AI

Many teachers and school leaders stressed the **need for professional support for teachers and school leaders on AI,** in the interviews. Some emphasised that AI was currenly an actively used technology, rather than a future innovation, and they highlighted the urgency of providing professional learning opportunities and ongoing support. They explained the need for a good understanding of the topic to answer students' questions, provide them with guidance and use AI wisely with them. Some also expressed their satisfaction that projects like AI4T were happening at this particular juncture as AI is becoming a prominent topic of discussion.

Our task is always to educate and obviously to do this, we ourselves must be trained. So, the big obstacle is teacher training. (Teacher, Italy)

Earlier, you talked about the potential dangers of AI. That's something I think is important to point out to students. But to talk about AI, you need to really understand the subject, and when I talk about a topic, I like to know it inside out and I'm not there with AI. It's not even that I'm not there, it's that I don't know what I'm talking about and I can't imagine bringing this up with students if I don't. (Teacher, France)

At the beginning of the project, we all asked ourselves: what is this all about? But a lot has happened since then. We have phenomenal timing here. The timing was really great. Just when this ChatGPT started. (Teacher, Slovenia)

The AI4T professional learning pathway has proved successful at improving teachers' knowledge of AI in all countries. Many teachers highlighted that they found the resources very informative and some commended the quality of the instructors' expertise. It would therefore be worth **using the learning resources produced during the AI4T project to address teachers' need to better understand AI**.

However, limitations were identified. In most countries, the AI4T professional learning pathway was often deemed **insufficient to support the use of AI in classroom practices.** For some teachers, the emphasis on theoretical aspects of AI (e.g. machine learning) was too substantial, leaving its practical aspects insufficiently explored. Indeed, teachers wanted their professional learning to be more focused on their **actual teaching needs**, including examples of use in specific course subjects and for specific grade levels. They also wanted this learning to facilitate professional practices, such as handling diverse student styles and differentiating their teaching. Therefore, **any future professional development pathway would accommodate teachers' expectations better if it were to contain moderate theoretical input and place a stronger emphasis on classroom practice.** This learning pathway could eventually become several modules, with one common, core module based on the AI4T resources, which would provide the basics for all teachers, and a second, more practical, module which would focus on teachers' specific needs within any given subject or level (e.g., AI tools for geometry in Year 10).

They didn't know right away how to apply AI in the classroom. Concretely, they didn't know the practical uses with the students. (School leader, France)

It seems to me that we teachers really value concrete experience, that is to say, from practice to practice. I mean, you know some applications, you know roughly what they do, but someone has explored them in depth, and it would make sense to say, 'Hey, look, with a linear function, you can do this and that'. GeoGebra, for example, I know it does a whole bunch of things, including programming and so on. But you might not have the energy to explore that. And maybe if someone showed you, you would at least try to use it. It would also be nice for students to show them, 'Look, using these basic principles in geometry, we can make a game in reality'. But you know, students ask all the time where I am going to use this. (Teacher, Slovenia)

I fully appreciate the idea of 'introducing' AI and so on, but at some point, it became too dense and too much information for me. I'm a regular teacher who is concerned with what I'm going to do today in class to reduce differences in levels between students, to encourage individualized pathways in order to meet actual student needs. (Teacher, France)

I found it was very interesting, but I found that there wasn't much towards the AI that we would be using in the class, and it was just, I suppose, it was really good to know about it, but I'd rather if there was a lot more, on what we would actually be doing in our classrooms. So I think a bit more focus on… how am I going to apply this in the classroom? What am I going to do? (Teacher, Ireland)

In the countries that provided blended learning, i.e. a combination of online and face-to-face sessions, many teachers expressed their satisfaction with this blended format. They mentioned the benefits, but also the limitations, of online resources, and they greatly appreciated the interactions during the face-to-face sessions. In Italy, the combination of online seminars and small-group webinars was also appreciated. This feedback suggests that **the use of the AI4T online learning resources should be incorporated in future professional development pathways as it allows for interactions and continuous support,** akin to what was offered in the AI4T project.

I think hybrid is always the way to go…it was good to have little groups in the meeting [face to face and in online seminars], different people from different schools, to hear what different people were doing, and then we could kind of bounce different ideas off each other. But also it's good that you can go online at 9pm and do a bit of work with the resources. So it's kind of the best of both worlds. (Teacher, Ireland)

Finally, several teachers stated that they would like to further their knowledge and **continue their professional learning of AI**. On this subject, some teachers and school leaders also emphasised the importance of **creating learning communities within schools**.

We think that artificial intelligence should be addressed in schools in a more transversal way, let's say not only with language and IT teachers, obviously, so we are trying to involve all teachers. (Teacher, Italy)

It would be great if this project could somehow be continued and expanded. To gradually add some teacher trainings, teacher seminars. ... But by subject area. For language teachers, for mathematics teachers. Or maybe natural sciences together, linguistics separately, humanities too. (Teacher, Slovenia)

School leaders would also benefit from professional development programmes on AI. As school leaders' self-assessed knowledge of AI was often reported as average (mostly "rather poor" and "rather good"), school leaders may benefit from resources that would facilitate their AI-related work (e.g., guidelines regarding the use of AI tools for education), and this, in turn, would facilitate teachers' work in the school. Further, at the organisational level, schools may consider communicating on AI-related matters with their digital teams.

[Interviewer: And do you think it would be relevant for school leaders to also receive training in AI?] Of course. Of course, it can be very relevant because we lack information on the subject. We can get some information from the media, but information that is specific and practical would be useful, necessary even. (School leader, France)

I personally believe that even our digital team is currently not, so to speak, trained and prepared on the topic... On the topic of artificial intelligence dedicated to teaching. Therefore, first, we must train the digital team. (School leader, Italy)

7.2 On the development of AI tools

Most teachers in the samples believed in the usefulness of AI for education. However, they wished they had more tools available to them and they pointed out the various obstacles restricting access to AI tools. For instance, many teaching tools require a fee. Generic AI tools such as ChatGPT created different problems, such as GDPR non-compliance. **Therefore, to support the use of AI by teachers, it is important to first ensure teachers' awareness of and access to AI tools that are free (or paid for in the long-term) and GDPR-compliant.**

I think there's a bit of disappointment regarding the tools. They thought they would have access to a bank of tools, a bank of applications. Well, things that could have been more readily usable with students and I think they got a few things out of it, but I think they were expecting more than that. (School Leader, France)

As I teach maths, we are given a site that is paid for [Kwyk]. Next year, if I want to use this site, I can't because I'd have to charge my students. So, I won't use it. (Teacher, France)

I'd wish there were more AI applications for the school in general (and for my subject/chemistry). (Teacher, Luxembourg)

Participants also stressed that **tools should answer specific professional needs**. These needs can vary depending on the course subject but also on the grade level and the types of students. Participants suggested that **including teachers, students and school leaders in the development of these tools,** in addition to implementing cost-benefit analyses of AI tools, would help address their needs more specifically.

There is a dilemma in developing platforms, different applications, programs, web robots and everything else. Is this really for education? I remember ... at a seminar in London, developers were saying, 'How wrong is it that we bring all this equipment and ask what part of it would be useful for you in education? We already have a product; you can't change it ... the screen is the way it is'. But maybe we should take the initiative so that the development comes from the field, from the practice. That might make sense. I mean, they probably have practitioners in their development departments. (School leader, Slovenia)

I think that if we want the tool to be used by teachers, it has to be conceptualised by teachers. In other words, I think that if a company or a group creates a tool by relying on artificial intelligence only without consulting with teachers, then there's little chance that the tool will be adapted to teachers' needs and used in the classroom. There definitely needs to be teachers on the conceptual development team or as product testers because it's quite complex. (School leader, France)

Maybe, how could it benefit the students and benefit the teachers, benefit their students' learning, like, see how their students could come on with it, and would it actually benefit them in a positive way… That maybe we need to teach the kids how to critically analyse what's coming back from these things, to make them more suitable… to adjust it to make it more pertinent to what they need (Teacher, Ireland)

In order to identify the professional needs that AI may be able to address, teachers answered an open question asking them to name superpowers that would help them in their work. Results showed that teachers wanted **help with marking students' papers**, **motivating or engaging learners**, **understanding students better, personalising their teaching,** and **finding and creating course materials. They also wished they had more time** in the day to do everything that they wanted to. These results match the categories which most teachers had identified in terms of the usefulness of AI, namely creating course content (88.9%), doing administrative tasks (87%), marking (81.5%), and monitoring students' work, learning progress or behaviour (84%). 57.6% of teachers also agreed that AI enables the personalisation of teaching to better meet students' needs. **Therefore, providing teachers with AI tools that address these specific tasks would prove particularly useful.**

[Marking] takes time, and it is interesting to understand student errors and to look at their levels of achievement. I think we need to continue doing this, but sometimes we might hold back on giving them tests, even formative ones, because it takes time to mark. So, I find tools like AI very interesting in our line of work, because they provide exercises, they grade them, and they give us an overview of students' levels of achievement. (Teacher, France)

We know very well that there are different academic levels in any student group, and that's even truer today than it was 20 years ago. So, teachers sometimes divide students based on their very, very different levels. We could imagine having very good students working on their own to a greater extent whilst the teacher spends a bit more time with students with difficulties using tools adapted to their needs. The issue here is that we can't let students get bored. [...] I think that, in any discipline, a tool such as an AI-based app would really contribute to managing students' academic differences in class. (School leader, France)

As a teacher I would kind of like the idea of wearing glasses for example, which would show me specific information about my students. How is there momentary mood or what is their attention level, in general or during a specific lesson...Or that I could quickly look up a student's average grade...Or let's imagine I had made a note regarding a specific student, which would be shown to me when looking at him/her with these glasses. (Teacher, Luxembourg)

Artificial intelligences serve to correct our mistakes as teachers. For example, a historical analysis of my assignments, the questions I ask and the errors can tell me if there are recurring errors on the same topic. Perhaps I expose that topic incorrectly or treat it poorly. I realise that this is a strong thing because for some of us it is very difficult to get off the desk. (Teacher, Italy)

I definitely see the future of school as a combination of these tools (technologies and AI) and aspects of traditional lessons. It is necessary to find a balance, managing to agree on everything. (Teacher, Italy)

7.3 On addressing ethical issues associated with AI

Most school leaders, teachers and students felt concerned with ethical issues associated with AI. Students were mostly concerned with the potential **loss of privacy** resulting from the collection of personal data by AI tools, but also with the **potential use of AI for illegitimate intents,** and with the difficulty of **attributing responsibility** when AI makes decisions for humans. Over half of the students responded that they were either "definitely" or "pretty much" concerned with these issues. AI transparency and any discrimination perpetuated by AI tools were seen as less of a concern. About half of the teachers thought that AI use in schools would **increase private companies' influence on schooling (52.9%)**, **surveillance in schools (49%), and place personal information at greater risk of being breached and used at people's expense (54.9%)**. Agreement on these issues was generally higher in France, Slovenia and Ireland than in Luxembourg and Italy.

Some teachers also highlighted issues of **equity in education,** emphasising the need for fair treatment of all students when assessing their work, ensuring that all students have access to AI tools, and checking the authenticity of student work, as well as considering the data being collected and shared.

> *Some things will have to be rethought... There is also the question of the access to these tools. It is a question of equity. The school system is supposed to be equitable, not only in Slovenia, but anywhere in the world ... Otherwise, it will be elitism all over again, and those who already have better opportunities will have even better opportunities in the future. (School leader, Slovenia)*

> *Where is the data going? What will we do with it? Is it stored? What use does it have? There are quite a few issues here. We work with students who are, let's face it, vulnerable, and who can be easily influenced, and so if we have bad intentions, we can also damage them and that's something that catches teachers' attention and they are very careful with this. So of course, there's a degree of mistrust. (School leader, France)*

> *AI is here; we cannot close our eyes. It will certainly provide some new opportunities for students to create their own shortcuts to grades in particular. So, it will probably be necessary to change the way education is delivered, to change the way assessments are made. For various written assignments, students can simply get help from ChatGPT. ... It might be necessary to start assessing in a more procedural way. (Teacher, Slovenia)*

> *So it's about the use of your time, you have that conundrum: OK it can really help with assessment, and can help learning, but you have to be really careful about the data that's being shared. So it's a balancing act. (Teacher, Ireland)*

In France, 62.5% of school leaders stated that they had not communicated with school staff over ethical guidelines regarding AI and data protection. Even though this figure is lower in Italy (38.2%) and in Slovenia (38.7%), **there is still room for improvement in terms of AI leadership, particularly in the realm of comprehensive communication strategies within schools**. Paradoxically, teachers and school **leaders emphasised that ethical concerns played an important role** in the integration of AI within schools, and they stressed the **need to supervise that integration on several levels.** Some participants mentioned the need for national authorities to **invest in the development of ethical AI tools for education**. Other participants requested that **AI tools be vetted by national institutions**. Putting forward **consistent guidelines on AI in schools** was deemed essential to many respondents.

> *We would absolutely need help from professional institutions so that we are not wandering around in the dark on our own, perhaps experimenting with the wrong things. The needs are already here, and now, we are already a bit late. (School leader, Slovenia)*

> *Just think of the investment! Everything that currently exists comes from either Asia or the United States. AI-ethical tools are being developed as we speak for the general public. But it's true that we would like the Education Nationale to get involved and take responsibility for these tools. (School Leader, France)*

> *We need to be presented with this tool and for it to be approved by inspectors to be sure that we stick to the rules, I think. (Teacher, France)*

> *Yeah, I think there should definitely be a set in stone AI policy for education. So there should be ... teachers shouldn't be worrying about what they're allowed or not allowed to do, using online tools. So I think there should definitely be a framework of the process, and what they can use these tools for. I think where people get frustrated about knowing what they are, what they are allowed to do, what they aren't allowed to do is when bad feelings start. (Teacher, Ireland)*

Conclusion

The evaluation focused on teachers' knowledge, perceptions and use of AI at the beginning of the experiment, then measured the impact of the intervention on these three outcomes, alongside considering additional factors such as the school context and students' attitude towards AI, which potentially influence teachers' outcomes.

At the beginning of the experiment, teachers had a moderate understanding of AI and displayed low familiarity with both technical terms and practical examples of AI tools. The impact analysis showed that the intervention was conducive to teachers' learning of AI in France, Italy and Slovenia. In France and Slovenia, teachers whose levels of engagement with the MOOC were higher also displayed greater learning gains. Although Ireland and Luxembourg were not taken into account for the impact analysis due to sample size, similar trends were observed in those samples, with teachers reporting a better level of knowledge after engaging with the professional learning pathway. In all countries, interviews with teachers also confirmed that they had learned more about AI thanks to the intervention. The success of the intervention in familiarising teachers with AI is evidenced by the overall high level of satisfaction with the different elements of the professional learning pathway. Open questions and interviews illustrated that teachers were satisfied with the various elements of the professional learning pathway because they were interesting and instructive. Overall, teachers appreciated the format consisting of online learning resources and face-to-face sessions or interactive webinars.

At the beginning of the experiment, most teachers had a positive view of artificial intelligence. They perceived AI as useful for their job, as enjoyable and as easy to use. On the other hand, few teachers reported anxiety when learning about or using AI. The intervention did not have a significant impact on teachers' perceptions of AI, except in Italy where perceived ease of use of AI increased for teachers in the intervention group. The lack of a significant effect of the intervention on perceptions may be linked to the pre-existing positive attitudes towards AI, leaving limited scope for enhancement. In addition, the project took place against a societal backdrop marked by technological progress in generative AI, and by the launch of ChatGPT during the experimental year. This launch received much divided media attention, which may have impacted public opinion on the topic of AI. Indeed, teachers had a slightly less positive view of AI by the end of the experiment in both the intervention and control groups. The broader societal context might have been a greater influence on teachers' perceptions than their participation in the AI4T project.

At the beginning of the experiment, over one-third of the teachers had not used educational AI tools, although some differences were noted across the countries: Slovenia, Ireland, and Luxembourg reported a higher level of use of AI tools, compared to France and Italy. Only in Italy did the intervention have a significant impact on the use of AI tools for all teachers who took part in the intervention. Results suggest that the intervention had a greater impact in Italy from a practical standpoint, demonstrated by an increase in both perceived ease of use of AI and actual use of AI. Practical differences in the professional learning pathway may explain the greater impact observed in Italy. For example, Italian teachers were organised into small groups, each being assigned a dedicated tutor. The school environment may also have played a role. While schools' ICT equipment and school leaders' support of the project were favourable elements present in all participating countries, a higher proportion of Italian schools had had prior involvement in AI-related projects. Furthermore, most school leaders in the Italian sample stated that integrating AI was a priority in their opinion – an opinion not as readily shared by the other countries. In France and Slovenia, the impact of the intervention on teachers' use of AI depended on their self-efficacy for integrating technology, and on their course subject. In France, an impact was observed on the use of two specific AI tools, Kwyk and Vittascience, and for teachers with higher self-efficacy, while in Slovenia, the overall use of AI increased for teachers with lower self-efficacy. Moreover, in these two countries, the impact on maths teachers' use of AI was significant, while the impact on language teachers' use of AI was close to 0. This result could be attributed to differences in AI tools made available to mathematics and language teachers. In France, for instance, maths teachers were given access to Kwyk, a tool tailored for mathematics, and they had the opportunity to test Vittascience, a coding tool, during the face-to-face session. However, despite mentions of ChatGPT and Duolingo during the intervention, language teachers reported being unable to use them, citing issues such as license fees or GDPR non-compliance. Finally, the impact on teachers' use of AI did not translate in the frequent use of the tools in any country, suggesting that the intervention led teachers to test tools rather than to fully integrate them into their teaching practice. This result is to be expected given the brief duration of the experiment and the professional learning pathway's emphasis on general AI understanding rather than on specific tool implementation guidance in classrooms.

Key insights from participants highlight the importance of offering professional development opportunities to teachers and school leaders in order to effectively support the integration of AI in education. The resources produced during the AI4T project proved valuable in familiarising teachers with AI. However, more comprehensive support is needed to enable them to integrate these tools in actual classroom practices. Furthermore, ensuring access to AI tools that address specific professional needs transpired to be a crucial factor. Finally, participants voiced apprehensions over ethical issues and the need for institutions to provide a safe environment for the use of AI. These concerns include data security and the potential misuse of AI for surveillance or corporate interests. These concerns were echoed by students, who expressed both awareness of and concern over data security and the use of AI for illegitimate intents.

AI4T
AI FOR TEACHERS

Appendices

Appendix A: Summary of the psychometric properties of the scales

Table 16: Summary of the psychometric properties of the scales for the teacher questionnaire

Name of the scale	Psychometric properties
Context	
Self-efficacy for integrating technology into the classroom	The scale includes 5 items. The Cronbach alpha is 0.93. The item-total correlations (Kendall's tau) are comprised between 0.76 and 0.78. There is one underlying factor that explains 72% of the variance. The factor loadings for each item are comprised between 0.84-0.86.
Reactions to the professional learning pathway	
Learner engagement	The scale includes 11 items. The Cronbach alpha is 0.86. The item-total correlations (Kendall's tau) are comprised between 0.41 and 0.55. There are four underlying factors. The first one explains 21% of the variance. On the first factor, the factor loadings for each item are comprised between 0.66-0.94. The second factor explains 18% of the variance. On the second factor, the factor loadings for each item are comprised between 0.69-0.79. The third factor explains 16% of the variance. On the third factor, the factor loadings for each item are comprised between 0.57-0.84. The fourth factor explains 14% of the variance. On the fourth factor, the factor loadings for each item are comprised between 0.69-0.94.
Satisfaction with the usefulness of the professional learning pathway	The scale includes 3 items. The Cronbach alpha is 0.92. The item-total correlations (Kendall's tau) are comprised between 0.82 and 0.87. There is one underlying factor that explains 79% of the variance. The factor loadings for each item are comprised between 0.85-0.94.
Participants' learning	
Knowledge of how AI works	The scale includes 5 items. The Cronbach alpha is 0.68. The item-total correlations (Kendall's tau) are comprised between 0.48 and 0.61. There is one underlying factor that explains 33% of the variance. The factor loadings for each item are comprised between 0.42-0.75.
Familiarity with AI technologies	The scale includes 5 items. The Cronbach alpha is 0.87. The item-total correlations (Kendall's tau) are comprised between 0.65 and 0.72. There is one underlying factor that explains 58% of the variance. The factor loadings for each item are comprised between 0.70-0.82.
Ability to identify AI tools	The scale includes 8 items. The Cronbach alpha is 0.77. The item-total correlations (Kendall's tau) are comprised between 0.31 and 0.59. There are two underlying factors. The first factor explains 31% of the variance. On the first factor, the factor loadings for each item are comprised between 0.67 and 0.89. The second factor explains 21% of the variance. On the second factor, the factor loadings for each item are comprised between 0.49 and 0.72.
Perceptions of AI	
Perceived ease of use of AI	The scale includes 4 items. The Cronbach alpha is 0.91. The item-total correlations (Kendall's tau) are comprised between 0.73 and 0.81. There is one underlying factor that explains 72% of the variance. The factor loadings for each item are comprised between 0.77-0.88.
Anxiety associated with use of AI and learning about AI	The scale includes 3 items. The Cronbach alpha is 0.90. The item-total correlations (Kendall's tau) are comprised between 0.74 and 0.83. There is one underlying factor that explains 69% of the variance. The factor loadings for each item are comprised between 0.73-0.91.
Enjoyment associated with use of AI and learning about AI	The scale includes 4 items. The Cronbach alpha is 0.90. The item-total correlations (Kendall's tau) are comprised between 0.79 and 0.85. The factor loadings for each item are comprised between 0.74-0.96.
Perceived usefulness of AI for education	The scale includes 10 items. The Cronbach alpha is 0.88. The item-total correlations (Kendall's tau) are comprised between 0.86 and 0.87. There is one underlying factor that explains 45% of the variance. The factor loadings for each item are comprised between 0.57-0.73.

Use of AI	
Use of AI	The scale includes 4 items. The Cronbach alpha is 0.9. The item-total correlations (Kendall's tau) are comprised between 0.79 and 0.82. There is one underlying factor that explains 69% of the variance. The factor loadings for each item are comprised between 0.77-0.88.
Frequent use of AI	The scale includes 4 items. The Cronbach alpha is 0.84. The item-total correlations (Kendall's tau) are comprised between 0.69 and 0.82. There is one underlying factor that explains 58% of the variance. The factor loadings for each item are comprised between 0.75-0.83.
Ethical consciousness when using AI	The scale includes 3 items. The Cronbach alpha is 0.75. The item-total correlations (Kendall's tau) are comprised between 0.70 and 0.76. There is one underlying factor that explains 56% of the variance. The factor loadings for each item are comprised between 0.53-0.94.
Intention to use AI	The scale includes 3 items. The Cronbach alpha is 0.88. The item-total correlations (Kendall's tau) are comprised between 0.82 and 0.86. There is one underlying factor that explains 74% of the variance. The factor loadings for each item are comprised between 0.69-0.95.

Table 17: Summary of the psychometric properties of the scales for the student questionnaire

Name of the scales	Psychometric properties
Attitude towards AI in education	The scale includes 8 items. The Cronbach alpha is 0.82. The item-total correlations (Kendall's tau) are comprised between 0.31 and 0.60. There are two underlying factors. The first factor explains 31% of the variance. On the first factor, the factor loadings for each item are comprised between 0.53 and 0.77. The second factor explains 12% of the variance. On the second fator, the factor loadings for each item are comprised between 0.53 and 0.64.
Concerns about ethical issues raised by AI in education	The scale includes 5 items. The Cronbach alpha is 0.82. The item-total correlations are comprised between 0.58 and 0.68. There is one underlying factor that explains 48% of the variance. The factor loadings are comprised between 0.61 and 0.75.

Appendix B: Comparisons of control variables and outcomes at the initial stage between the control group and the intervention group

Table 18: Comparisons of control variables in the intervention and control groups at the beginning of the experiment

Control variable	Control group	Intervention group	p-value
France			
Sex	0,49	0,37	**0,02****
Teaching experience *(Average number of years of teaching experience)*	18,26	18,94	0,45
Class size *(Number of students in the class participating in the experiment)*	27,00	27,38	0,55
Student academic difficulties *(Percentage of students with academic difficulties in the class)*	37,57	33,11	**0,01*****
Italy			
Sex	0,30	0,28	0,63
Teaching experience *(Average number of years of teaching experience)*	17,20	20,10	**0,01*****
Class size *(Number of students in the class participating in the experiment)*	20,92	20,42	0,17
Student academic difficulties *(Percentage of students with academic difficulties in the class)*	28,30	24,93	**0,01*****
Slovenia			
Sex	0,14	0,28	**0,01*****
Teaching experience *(Average number of years of teaching experience)*	18,17	15,27	**0,01*****
Class size *(Number of students in the class participating in the experiment)*	24,93	23,87	**0,07****
Student academic difficulties *(Percentage of students with academic difficulties in the class)*	43,21	43,40	0,94

Table 19: Comparisons of the mean in outcomes related to knowledge of AI between the intervention and control groups at the beginning of the experiment

Outcome	Control group	Intervention group	p-value
France			
Self-assessment of knowledge	-0,01	0,00	0,66
Knowledge of how AI works	-0,16	0,00	0,31
Familiarity with AI technologies	-0,03	0,00	0,83
Ability to identify tools mainly based on AI as AI	-0,45	0,00	**0,01*****
Ability to identify tools that are are not mainly based on AI as AI	-0,21	0,00	0,13
Slovenia			
Self-assessment of knowledge	-0,01	0,00	0,94
Knowledge of how AI works	0,22	0,00	**0,07***
Familiarity with AI technologies	0,04	0,08	0,71
Ability to identify tools mainly based on AI as AI	-0,02	0,00	0,85

Ability to identify tools that are are not mainly based on AI as AI	-0,06	0,00	0,58
Italy			
Self-assessment of knowledge	-0,09	0,00	0,45
Knowledge of how AI works	-0,17	0,01	0,17
Familiarity with AI technologies	0,02	0,09	0,46
Ability to identify tools mainly based on AI as AI	-0,11	0,00	0,40
Ability to identify tools that are are not mainly based on AI as AI	0,10	0,00	0,42

Table 20: Comparisons of the means in outcomes related to perceptions of AI between the intervention and control groups at the beginning of the experiment

Outcome	Control group	Intervention group	p-value
France			
Perceived ease of use of AI	0,04	0,00	0,83
AI anxiety	0,00	0,00	0,97
AI enjoyment	-0,02	0,00	0,93
Perceived usefulness of AI for teaching	-0,25	0,00	0,11
Slovenia			
Perceived ease of use of AI	0,03	0,00	0,85
AI anxiety	-0,10	0,00	0,41
AI enjoyment	0,02	0,00	0,91
Perceived utility of AI for teaching	0,02	0,00	0,81
Italy			
Perceived ease of use of AI	-0,09	0,00	0,44
AI anxiety	0,11	0,00	0,39
AI enjoyment	-0,13	0,00	0,27
Perceived utility of AI for teaching	-0,16	0,01	0,20

Table 21: Comparisons of the means in outcomes related to use of AI between the intervention and control groups at the beginning of the experiment

Outcome	Control group	Intervention group	p-value
France			
Use of AI tools	-0,14	0,00	0,36
Frequent use of AI tools (at least once a week)	-0,12	0,00	0,40
Ethical consciousness when using AI	0,12	0,01	0,57
Intention to use AI tools	-0,39	0,00	**0,01*** **
Slovenia			
Use of AI tools	0,05	0,00	0,70
Frequent use of AI tools (at least once a week)	0,08	0,00	0,51
Ethical consciousness when using AI	-0,14	-0,02	0,37
Intention to use AI tools	0,10	0,00	0,39
Italy			
Use of AI tools	0,11	0,00	0,39
Frequent use of AI tools (at least once a week)	0,00	0,00	0,99
Ethical consciousness when using AI	-0,06	-0,03	0,84
Intention to use AI tools	-0,03	0,00	0,80

References

Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 1- 10.

Anderson, R. E., & Dexter, S. (2005). School Technology Leadership: An Empirical Investigation of Prevalence and Effect. *Educational Administration Quarterly*, *41*(1), 49- 82. https://doi.org/10.1177/0013161X04269517

Badia, Antoni, Julio Meneses, Carles Sigalés, et Sergi Fàbregues. « Factors Affecting School Teachers' Perceptions of the Instructional Benefits of Digital Technology ». *Procedia - Social and Behavioral Sciences* 141 (août 2014): 35762. https://doi.org/10.1016/j.sbspro.2014.05.063.

Banerjee, A. V., & Duflo, E. (2017). An introduction to the "Handbook of Field Experiments." *Handbook of economic field experiments*, *1*, 1- 24.

Casper, Wm, Bryan Edwards, Craig Wallace, Ronald Landis, et Dustin Fife. « Selecting response anchors with equal intervals for summated rating scales ». *Journal of Applied Psychology* 105 (15 août 2019). https://doi.org/10.1037/apl0000444.

Cave, S., Coughlan, K., & Dihal, K. (2019). *'Scary Robots': Examining public responses to AI*. https://doi.org/10.17863/CAM.35741

Céci, Jean-François. « Analyse des pratiques numériques des enseignants, du collège à l'université, au prisme du genre ». *IJARTech (International Journal of Applied Research and Technology)*, Articles JIP2018, 1 (janvier 2019). https://hal.archives-ouvertes.fr/hal-01994895.

Christensen, R. W., & Knezek, G. A. (2009). Construct validity for the teachers' attitudes toward computers questionnaire. *Journal of computing in Teacher Education*, *25*(4), 143- 155.

Commission Européenne. (2022). *Plan d'action en matière d'éducation numérique (2021-2027)*. https://education.ec.europa.eu/fr/focus-topics/digital-education/action-plan

Davis, F., Bagozzi, R., & Warshaw, P. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*, 982- 1003. https://doi.org/10.1287/mnsc.35.8.982

Deng, R., Benckendorff, P., & Gannaway, D. (2020a). Learner engagement in MOOCs: Scale development and validation. *British Journal of Educational Technology*, *51*(1), 245- 262. https://doi.org/10.1111/bjet.12810

Deng, R., Benckendorff, P., & Gannaway, D. (2020b). Linking learner factors, teaching context, and engagement patterns with MOOC learning outcomes. *Journal of Computer Assisted Learning*, *36*(5), 688- 708. https://doi.org/10.1111/jcal.12437

European Commission, Directorate-General for Education, Youth. *Final Report of the Commission Expert Group on Artificial Intelligence and Data in Education and Training: A Executive Summary*. LU: Publications Office of the European Union, 2022. https://data.europa.eu/doi/10.2766/65087.

« A Definition of AI: Main Capabilities and Scientific Disciplines ». European Commission, High-Level Expert Group on Artificial Intelligenc, 2019. https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines.

Ertmer, Peggy A., Anne T. Ottenbreit-Leftwich, Olgun Sadik, Emine Sendurur, et Polat Sendurur. « Teacher Beliefs and Technology Integration Practices: A Critical Relationship ». *Computers & Education* 59, nº 2 (1 septembre 2012): 42335. https://doi.org/10.1016/j.compedu.2012.02.001.

Fengchun, Miao, Wayne Holmes, Huang Ronghuai, et Zhang Hui. « AI and education: guidance for policy-makers ». UNESCO, 2021. https://unesdoc.unesco.org/ark:/48223/pf0000376709.

Février, F., Gauducheau, N., Jamet, É., Rouxel, G., & Salembier, P. (2011). The study of affects in human-computer interactions: Theories, methods and benefits. *Le travail humain*, *74*(2), 183- 201.

Guskey, T. R. (2000). *Evaluating professional development*. Corwin press.

Jang, Y., Choi, S., & Kim, H. (2022). *Development and validation of an instrument to measure undergraduate students' attitudes toward the ethics of artificial intelligence (AT-EAI) and analysis of its difference by gender and experience of AI education*. https://link.springer.com/article/10.1007/s10639-022-11086-5

Njiku, J., Maniraho, J. F., & Mutarutinya, V. (2019). Understanding teachers' attitude towards computer technology integration in education: A review of literature. *Education and Information Technologies*, *24*(5), 3041- 3052. https://doi.org/10.1007/s10639-019-09917-z

Noiwan, J., Piyawat, T., & Norcio, A. F. (2005). *Computer Attitude and Computer Self-Efficacy: A Case Study of Thai Undergraduate Students*. 11.

Perrotta, Carlo. « Do School-Level Factors Influence the Educational Benefits of Digital Technology? A Critical Analysis of Teachers' Perceptions: The Educational Benefits of Digital Technology Use ». *British Journal of Educational Technology* 44, nº 2 (mars 2013): 31427. https://doi.org/10.1111/j.1467-8535.2012.01304.x.

Poyet, Françoise. « Perception de l'utilité et usages pédagogiques d'environnements numériques de travail par des enseignants du second degré ». *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation* 22, nº 1 (2015): 4564. https://doi.org/10.3406/stice.2015.1686.

Remian, D. (2019). Augmenting Education: Ethical Considerations for Incorporating Artificial Intelligence in Education. *Instructional Design Capstones Collection*. https://scholarworks.umb.edu/instruction_capstone/52

Samoili, Sofia, COBO Montserrat Lopez, Blagoj Delipetrev, Fernando Martinez-Plumed, GUTIERREZ Emilia Gomez, et PRATO Giuditta De. « AI Watch. Defining Artificial Intelligence 2.0 », 29 octobre 2021. https://doi.org/10.2760/019901.

Schepman, A., & Rodway, P. (2020). Initial validation of the general attitudes towards Artificial Intelligence Scale. *Computers in Human Behavior Reports*, *1*, 100014. https://doi.org/10.1016/j.chbr.2020.100014

Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, *128*, 13- 35.

Schiff, D. (2021). Out of the laboratory and into the classroom: The future of artificial intelligence in education. *AI & SOCIETY*, *36*(1), 331- 348. https://doi.org/10.1007/s00146-020-01033-8

Shattuck, G. (2009). Understanding School Leaders' Role in Teachers' Adoption of Technology Integration Classroom Practices. *Educational Media and Technology Yearbook*, 7- 28.

Suh, W., & Ahn, S. (2022). Development and Validation of a Scale Measuring Student Attitudes Toward Artificial Intelligence. *SAGE Open*, *12*, 215824402211004. https://doi.org/10.1177/21582440221100463

Tricot, A. (2020). [Report] Quelles fonctions pédagogiques bénéficient des apports du numérique. Numérique et apprentissages scolaires. https://ecogestion-caen.second-degre.ac-normandie.fr/IMG/pdf/201015_cnesco_tricot_numerique_fonctions_pedagogiques-1.pdf

Wang, B., Rau, P.-L. P., & Yuan, T. (2022). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technolog*. https://www.tandfonline.com/doi/abs/10.1080/0144929X.2022.2072768?journalCode=tbit20

Wang, Y.-Y., & Wang, Y.-S. (2019). Development and validation of an artificial intelligence anxiety scale: An initial application in predicting motivated learning behavior. *Interactive Learning Environments*, *0*(0), 1- 16. https://doi.org/10.1080/10494820.2019.1674887

Yennek, N. (2014). Contribution de l'intérêt situationnel à une reconsidération de la satisfaction dans la formation pour adultes [Thesis, Theses.fr]. In *Theses.fr*. https://www.theses.fr/2014PA100122