1. AI in Education

1.1 Potentials & Risks

> *If you remember only one sentence form this page:* AI has made its way into schools like a medicine that comes without a patient information leaflet. As a teacher, you don’t know when to use AI, you lack dosage instructions, and there is no information on the side effects. This chapter serves as the patient information leaflet for the Playbook.

Should AI come with a leaflet like a medicine?

You know that slip of paper tucked inside every medicine box? The one you likely toss aside without a glance? It may seem dull, but this little paper is a quiet triumph of the last century. That leaflet spells out four essentials: what the medicine treats, how much to take, the surprises it might bring, and who should steer clear.

Behind every line on that leaflet stands a whole machinery: clinical trials, regulators, and a responsibility to reveal mistakes. No one believes the leaflet itself heals you. Its real power is making the medicine safe for anyone. That is the magic we are missing when we let AI teach without guidance.

The AI tools now quietly settling into your school arrived with no instructions. In the UK, a review found that only 7 out of 100 education technology companies had ever put their products through a controlled trial, and only 12 out of 100 had sought outside certification. The classroom tells the same story. Only 11 out of 100 school leaders and teachers across 17 American states had ever requested peer-reviewed proof before adopting a new tool.

The result, in many cases, is wasted money. In the US, two-thirds of school software licenses gathered dust, and 98% saw almost no use. The EdTech Genome Project looked at about 7,000 teaching tools worth 13 billion dollars and found that 85% were either mismatched or misused.

This chapter may look like the dry, necessary section, but it reveals the hidden stories behind the rules. It tells what really happened, who was affected, and shines a light on the side effects, even those that quietly reached children who never signed up.

Potentials of AI

Let's make this clear right from the start. Most studies on AI's potential focus on university students. School pupils have so far featured relatively rarely in these studies. Most studies relate to

short-term interventions lasting from a few hours to a few weeks. Consequently, the so-called novelty effect may have positively influenced the results. Last but not least, the test used to assess effectiveness was often devised by the very people who delivered the lessons. This, too, may have confounded the data.

(1) Let AI explain a new topic and keep the practice for your lesson

Imagine this. Tomorrow, when a pupil misses a topic or needs that tricky concept explained yet again, let an AI tutor handle the first explanation. Reserve the practice for your lesson, where real learning comes alive.

Why this matters. AI can support acquisition of factual knowledge. In a Harvard study, 194 physics students learned two topics. For one topic, they used an AI tutor developed for this purpose. For the other topic, they learned in a traditional class. Each student completed both approaches, so no one can claim that the stronger group received the better treatment. In a test conducted immediately afterward, the students taught by the AI achieved significantly higher results. The researchers recorded a median time of 49 minutes to complete the task, compared with the 60 minutes allocated for the lesson in the timetable. The authors are cautious in their interpretation of these results: the material was new to the students, and much of the task simply involved explaining it well. They expressly reject the idea that the same would apply to tasks requiring the integration of multiple ideas. Furthermore, these were undergraduate students, not Year 8 pupils.

A large review pooling the results of 228 studies came to a similar conclusion. Generative AI-based chatbots produced a larger effect on knowing (and understanding) than intelligent tutoring systems and adaptive practice software.

Go deeper. Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports, 15,* Article 17458. Free access, DOI 10.1038/s41598-025-97652-6. The article describes how the tutor was designed and controlled. This is useful for anyone wishing to develop an AI tutor themselves.

(2) Making better use of lesson planning time

Imagine this. Save some of the time you need for lesson planning by using AI to draft your first lesson plan and then revise it thoroughly.

Why this matters. No one disputes that teachers' workloads are heavy. Therefore, one potential benefit of AI is using working hours more efficiently. A UK study investigated the extent to which the targeted use of AI for lesson planning can help to reduce teachers' workload. Researchers randomly assigned 259 science teachers at 68 schools to either a group using a chatbot alongside brief written instructions or a group that continued lesson planning as before. **The weekly planning time for science lessons in Years 7 and 8 fell from around 81 minutes to around 56.** A panel of subject experts assessed the quality of the lesson plans. They did not know whether the plans had been created with or without the aid of AI. The experts found no evidence to suggest that the quality of the lesson plans differed. However, two points must be highlighted.

(1) **The teachers used the tool less and less as the weeks went on,** meaning that the time saved does not occur automatically and does not sustain itself. (2) The trial measured the time spent by teachers on lesson planning, not their pupils' learning outcomes.

Go deeper. Roy, P., Poet, H., Staunton, R., Aston, K., & Thomas, D. (2024, 12 December). *ChatGPT in lesson preparation: A Teacher Choices Trial.* NFER, commissioned by the Education Endowment Foundation and the Hg Foundation. Free at educationendowmentfoundation.org.uk.

(3) Providing better support for learners who do not benefit fully from standard lessons

Imagine this. Your school decides how it wishes to spend its funds. It wants to purchase an AI chatbot. How likely do you think it is that the chatbot providers will provide you with evidence of the bot's benefits for different subjects?

Why this matters. The provider may point out the following application. Chatbots are the oldest and best-documented application of 'AI' in education, and this is precisely what is often overlooked when allocating funds. In an analysis of 29 studies involving 41 groups of learners with disabilities, researchers found moderate benefits across a wide range of settings. But read the small print. In **seven out of ten of these studies**, 'AI' meant a **robot**: a small physical machine sitting on a table, often operated by a researcher in another room. Only one in five studies used software. So this is no proof that a chatbot will help an autistic pupil in your class. Nor could the reviewers say **what** makes these interventions effective. Based on their research, they could not identify why the studies that worked differed from those that did not.

Go deeper. Zhang, L., Carter, R. A., Jr., Liu, Y., & Peng, P. (2026). Let's CHAT about artificial intelligence for students with disabilities: A systematic literature review and meta-analysis. *Review of Educational Research, 96* (1). DOI 10.3102/00346543241293424. The real gem is the table of included studies: it instantly reveals just how little of this field actually focuses on the tools currently being marketed to schools.

(4) Teaching learners how AI fails

Imagine this. Ask your class to create ten images of 'a doctor' and ten of 'a nurse.' Tally up who appears in each set, then spark a discussion: where did these patterns come from, and who made those choices?

Fig 1. Image created with NanoBanana 2, created August, 20, 2026, One shot prompt: a doctor, a nurse

Why this matters. The classic version of this lesson feels outdated, and that is exactly why it deserves attention. For years, the demonstration was predictable: ask for a doctor, and a man appears. In one systematic test, Midjourney almost always pictured doctors as white men, while Adobe Firefly made visible, if imperfect, attempts to diversify.

Try the same prompt today, and you might see a female doctor or a male nurse appear (Fig. 1). But do not be fooled into thinking the bias has vanished. A 2026 analysis of 1,344 images from three popular generators revealed the opposite: when asked for a 'competent person,' one system showed no women.

What changed is not the bias but the layer it lives in - and this the important lesson! Human decisions stand behind any of these pictures: what the system was trained on, and what its makers afterward decided it should show. The second is adjustable, invisible, and made by a company.

The first layer of human decisions is what the system has absorbed. Millions of images, crafted and labeled by people, each one echoing the world as it already was. This is the layer that the classic bias lesson reveals, and pupils grasp it easily: the machine became a reflection of what it was fed.

The second layer is shaped by what a company chooses to show you next, yet this influence leaves no trace in the image itself. Three things unfold here, each one described by the creators in their own words.

- **Your words are rewritten before the model receives them.** From OpenAI's own developer documentation: "we use GPT-4 to optimize all of your prompts before they're passed to DALL-E" — and "this feature isn't able to be disabled at the moment." What your pupil typed is not what the machine was asked.
- **Appearance gets filled in when the prompt leaves it open.** OpenAI, describing its diversity technique: "This technique is applied at the system level when DALL·E is given a prompt describing a person that does not specify race or gender, like 'firefighter.'" Afterwards, users were twelve times more likely to say the images showed people of diverse backgrounds. "A doctor" is exactly that kind of prompt.
- **And that tuning can miss.** Google, after pausing Gemini's image generation in 2024: "Our tuning to ensure that Gemini showed a range of people failed to account for cases that should clearly not show a range."

So when a pupil types "a doctor" and a woman appears, the image offers no clue about which layer brought her into being. **That uncertainty is not a flaw in the lesson. It is the lesson itself.** The old story claimed the machine simply soaked up society's prejudices. The new story cuts deeper and lasts longer: the image is no mirror, and it is no accident. A company made a choice! You cannot spot that choice in the outcome, and it might shift on a Tuesday, quietly, without warning. That truth endures through every new model. The stereotype lesson fades away.

This is the lesson that deserves a student's attention, and it will outlast the next model update: **the image is not a simple reflection of reality, and it did not happen by chance. Someone made a choice.**

This lesson can start early. When 209 Finnish students in 12 classes, grades 4 and 7, explored this topic, the number who could explain the bias using data jumped from about 7 out of 100 to 44 out of 100. The honest reservation: the study measured what children could explain, not what they would do later.

Go deeper. Vartiainen, H., et al (2025). Enhancing children's understanding of algorithmic biases in and with text-to-image generative AI. *New Media & Society, 27*(9). Free, DOI 10.1177/14614448241252820

Weinmann, H., et al (2026). Gender bias in text-to-image generative artificial intelligence: Neglect and stereotypical presentations across three popular platforms. *New Media & Society,* Online First

How solid are the research results on the potentials of AI at this very moment?

The most important thing. If someone quotes a study to you, ask two questions before agreeing: "Who were the learners, and how long did the study last?"

Why this matters. In 2025, a meta-analysis of 51 studies reported a **significant benefit** of ChatGPT for pupils' academic performance. The article was read hundreds of thousands of times

and served as the basis for numerous confident statements in staff rooms. **On 22 April 2026, the journal retracted the article.** A *retraction* means that a journal officially withdraws something it has published; the article remains visible, marked with the note ‘RETRACTED’, so that anyone who has cited it can recognise this. Two external researchers had identified inconsistencies in the synthesis of the studies. The journal's editor wrote that the problems ‘undermine the editor’s confidence in the validity of the analysis’. The authors did not respond to correspondence on this matter.

This does not mean that AI has no impact. There are reputable studies that demonstrate moderate benefits. This field of research is also not yet well defined because it is relatively new and highly dynamic. It means that the figure that was frequently cited was incorrect, and that it took a year and two external experts to realise this.

Another reservation worth mentioning is: The closer the evidence gets to school, at least at the moment, the smaller the effect.

Go deeper. The retraction itself is one page, free, and worth reading precisely because it is short: Wang, J., & Fan, W. (2026). Retraction note: The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking. *Humanities and Social Sciences Communications, 13*, 528. DOI 10.1057/s41599-026-07310-z.

The **retracted** original was *HSSC 12,* 621 (2025). Still standing: Wu, X., Zhu, P., Zhang, J., Yin, M., & Wang, Y. (2026). *HSSC, 13*, 684. Free, DOI 10.1057/s41599-026-07019-z.
The school-only figure: Yi, L., Liu, D., Jiang, T., & Xian, Y. (2025). *International Journal of Science and Mathematics Education, 23*(4), 1105–1126. DOI 10.1007/s10763-024-10499-7.

Risks of AI

(1) Doing something is not the same as learning

Imagine this. Before you set a homework task that a chatbot can complete, think about what the purpose of the task is. Might it be a good idea to change the task?

Why this matters. Do you learn faster with a chatbot? Around a thousand pupils in Years 9 to 11 practised maths in three groups: one using a standard chatbot, one using a version programmed to withhold the answer and ask a follow-up question, and one without a chatbot. Whilst using the chatbot, the group with the standard chatbot completed 48 percent more practice exercises than the group without one. Afterwards, everyone sat the same exam without a chatbot. The result, which seemed surprising at first, was that the group using the standard chatbot performed 17 percent worse than the pupils who had not used a chatbot. The group that had worked with the chatbot which withheld the answer showed no drop in performance in the test compared to the group without a chatbot, but practised more. The pupils using the standard chatbot were neither lazy nor did they cheat. They worked harder and learnt less, because what the tool eliminated was not the effort, but the cognitive challenge that constitutes learning.

Go deeper. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *PNAS, 122*, e2422633122. DOI 10.1073/pnas.2422633122.

(2) AI Detectors punish the wrong pupils

Imagine this. You receive a student's essay to grade. Curious about its authenticity, you feed it into an AI detector. The result flashes: the essay was almost certainly written by AI.

Why it matters. You use the AI detector to ensure fairness, **yet for a particular group, the detectors have exactly the opposite effect.** A detector does not recognise machine-generated text. It measures how predictable the wording is. And that is precisely where the problem lies, for example, when a pupil is writing in a foreign language, is unpractised in writing, or has a limited everyday vocabulary. Their writing is predictable. When seven commercial detectors analysed 91 essays written by non-native speakers sitting an English exam, these essays were incorrectly flagged as machine-generated in around six out of ten cases. By contrast, essays written by 14-year-old Americans in their native language, English, were almost all classified correctly. A tool that triggers false positives for exactly one group is not neutral. Furthermore, this group is the least able to defend itself.

Go deeper. Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns, 4*(7), 100779. Free, DOI 10.1016/j.patter.2023.100779.

(3) Where on earth will the pupils' data be?

Imagine this. You use an AI tool your school offers. Try to find out where the pupils' data go.

Why it matters. Can your school ensure that the data is not passed on to a third party? A third party is company that is neither your school nor the app's developer. Typically, these are advertising or analytics companies that receive data whilst the pupil is working. No one at the school chose this company. When researchers carried out a technical analysis of 163 learning products recommended by governments during the pandemic-related school closures, 145 of them processed children's data in a way that jeopardised or infringed their rights by transmitting data to 196 third-party company or granting them access, predominantly from the advertising industry. **Of the 42 governments that developed their own products rather than purchasing them, 39 had developed products with the same problem. The study covers products from 2021, and some may have changed since then**. However, we cannot rule out that the same applies to AI applications.

Go deeper. Human Rights Watch (2022). "*How dare they peep into my private life?:* Children's rights violations by governments that endorsed online learning during the Covid-19 pandemic. Free at hrw.org. The country annexes let you look up what was endorsed where.

(4) AI arrives last where it is needed most

Imagine this. How much training did you already receive on learning about AI or subject-specific teaching with AI?

Why this matters. Adults who had completed their education were provided with an AI assistant. The performance gap between those with higher and lower levels of formal education narrowed significantly. When the researchers then withdrew the assistant, the gap partially re-emerged. The researchers concluded that the AI assistant had carried out the work, but had not helped participants develop any skills.

A survey of American school principals revealed the following picture: the more disadvantaged a school's pupils were, the less impact AI had on teaching.

If we combine the findings from both studies, they allow us to make the following statement: A good tool can bridge a gap as long as it is in the hands of a learner. However, it usually does not reach those who need it, and leaves no skills behind when it is removed.

German figures show the same patte. Among 1,590 young people aged 14 to 20, 80 percent in the most affluent families and 55 percent in the least affluent saw AI as an opportunity. And recognition is followed by action: 70 percent of grammar school pupils used AI for homework, compared with 58 percent at lower secondary schools.

What about teacher professional development? A survey of more than ten thousand teachers in England found that 45 percent of teachers at private schools had already received training on AI. This compared with 21 percent at state schools. Within the state school system, the figures ranged from 26 percent at the wealthiest schools to 18 percent at the poorest schools.

(5) What improves when learning with AI is not necessarily what you see

Imagine this. You introduce a new ChatBot into your teaching and the pupils enthusiastic about using the new tool.

Why this matters. A large review found four different things about learning (Knowing and understanding, working with AI, engagement, motivation, confidence, ability to plan, monitor and correct their own work). The last one didn't improve whereas the others did.

- For knowing and understanding generative AI had larger effects than intelligent tutoring systems and adaptive practice software that is already used in schools.
- For applying knowledge, and for motivation and engagement, AI did not perform better that already existing software.
- For self-steering we do not have enough studies to tell any effect.

Here's a quick side note: t**he greatest impacts are found in the arts and humanities, rather than in mathematics or the natural sciences.**

Go deeper. Yeo, G. H., & Lansford, J. E. (2025). Effects of artificial intelligence on educational functioning: A review and meta-analysis. *Educational Psychology Review, 37*(4), article 110. DOI 10.1007/s10648-025-10085-5.

1.2 Teaching About vs Teaching with AI

Take Home message

- ***AI is a satnav.*** AI can save the driver who already carries the map and prevent the one who doesn't from learning. The effect depends on the person, not the product.
- ***Teaching about AI is the precondition for Teaching with AI. Y***ou need to know the machine to check the answer the machine produced.
- **Knowing is not doing.** Teachers rate their AI knowledge above their ability to teach with it, and training alone does not close that gap.
- ***Teachers use AI backstage, not in class.*** Half of the teachers use AI regularly, under a third use it with students.
- ***Pedagogy beats technology.*** Transforming lessons beats chatting with AI; working in pairs beats working alone; feedback beats gamification.
- ***Better feedback is not better learning.*** AI comments judged clearly better than teachers' produced revisions that were no better at all. What decides is whether feedback is *taken up.*
- ***AI raises output, not self-directed learning.***
- ***Clear rules and heavy use go together.*** Teachers in schools with a written AI policy report saving more time, not less — the opposite of what most people expect.
- ***Guard your own map.*** *Relief becomes deskilling where AI absorbs the work that builds and maintains judgment.*

> *If your remember only one sentence from this page:* **AI is a satnav. It saves the driver who already has a map in her head and quietly stops the driver who hasn't from ever building one**

The satnav problem

To drive a London taxi, you must first pass a knowledge test: 320 set routes across the 113 square miles within six miles of Charing Cross, for which Transport for London allows up to two years before the examinations even begin.

Examined drivers also use a satnav, and nobody thinks less of them for it. The knowledge test did not become worthless when the device arrived. The satnav became the thing that makes the drive safe. A driver with the city in her head uses the satnav for what she cannot know: the accident on the bridge or tonight's traffic. And she notices when it routes her down a street that has been dug up since April.

Now put a newcomer in the same taxi with the same satnav. She arrives at every address. For a while, she is indistinguishable from the experienced driver. But following the satnav's instructions, she is building nothing. After a year of turning left when instructed, she still cannot cross the city without the device, and she cannot tell when it is wrong.

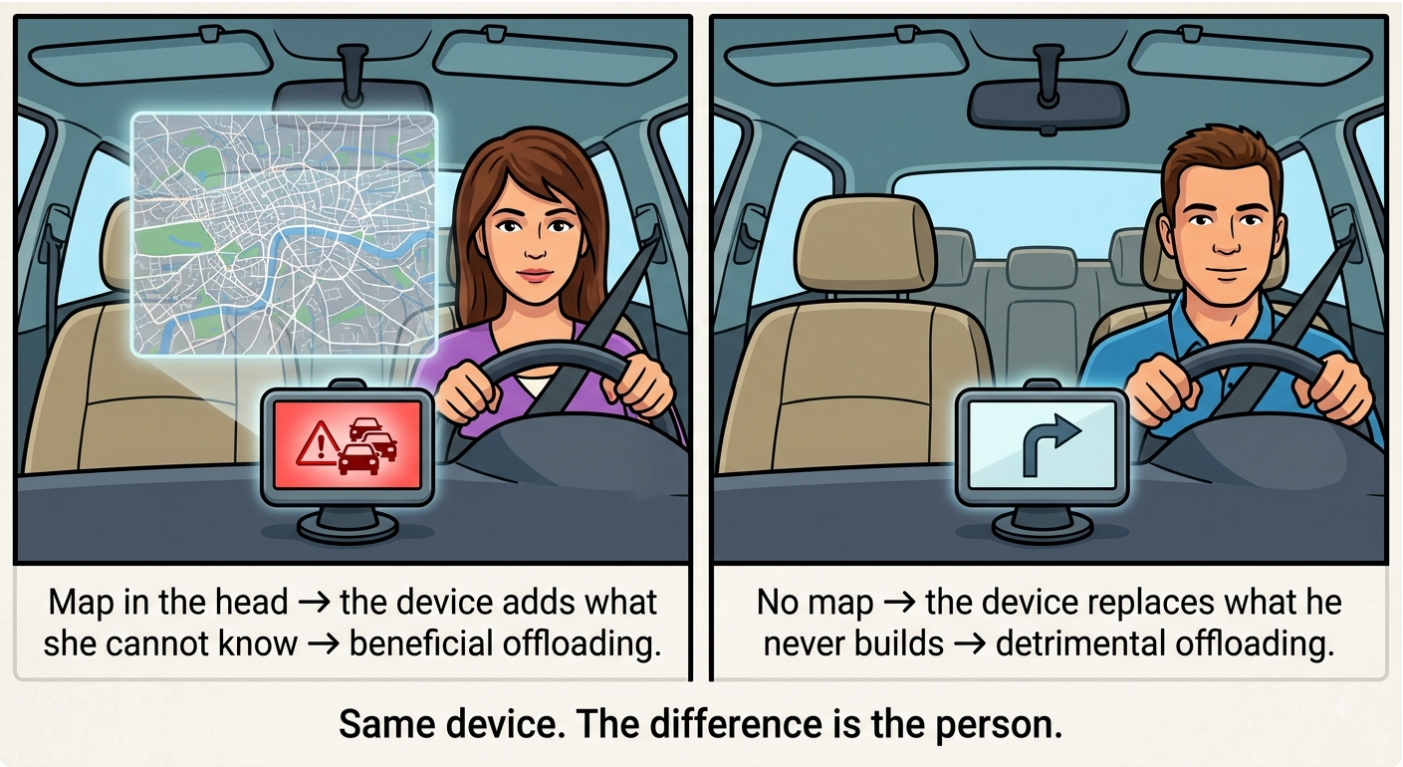

Fig. 1. The outcome depends on the driver. Image created with GoogleGemini.

Same satnav, two drivers, opposite outcomes. **The difference was never in the device.**

That is a decade of research on AI use in one image, and it holds without the image too. The effect of an AI tool depends far more on what the user already knows than on the tool itself. This is what researchers call a schema, the organized prior knowledge of a person. Where it is in place, handing work over to a machine frees you up. Researchers call this beneficial cognitive offloading. However, when knowledge needs to be developed, the same process can prevent it from forming. This is called detrimental cognitive offloading.

A study with around a thousand secondary students shows this happening in a mathematics classroom. One group practiced with an unrestricted chatbot, another without. While they had it, the chatbot group solved **48 % more practice problems!** Then came an exam with the chatbot taken away. Now, the same students who solved practice problems so well scored **17% worse** than the group that had never used the AI chatbot. A third group used an AI version built to withhold answers and ask questions instead. That group lost nothing.

That third group is the whole design question. A satnav that names the destination and helps you find the route teaches you the city. A satnav that says *turn left in 200 meters* does not, however good its data. Researchers call the difference guardrails, and the **design decides whether you help your students learn the city's map or simply find the destination.**

Please remember: **You have the knowledge. Your students, in the subject you teach, do not yet.** That asymmetry is the whole chapter.

Two floors, not two topics

In academic literature, these two areas are considered distinct fields. You could think of them as two floors of a building, with most teachers located on the ground floor and no staircase leading to the first floor.

When asked to rate their knowledge of AI, teachers at German vocational schools gave themselves an average rating of 3.4. However, they rated their ability to apply this knowledge in the classroom much lower, at 2.6. Knowing and doing had become disconnected. A trial across five European countries then tested whether training could close that gap. In the trial, 736 teachers in France, Ireland, Italy, Luxembourg and Slovenia were given a course and support. The results showed that their knowledge and ability to judge what a tool could do increased. However, their teaching did not change. They tried things and went back to what they had been doing before.

This is the knowledge–action gap. Courses produce knowledge, but changing practice needs accompaniment, trialing and colleagues to think with. **Transfer has to be designed, not hoped for.**

What teachers are actually doing

The most robust finding across countries is quietly surprising: t**eachers use AI intensively for themselves and hardly at all with their students**. In Germany, 52% use AI at least once a month. Only 29 % use it in lessons with a class in front of them. So the public debate about "AI in the classroom" is largely a debate about something that is not yet happening at scale.

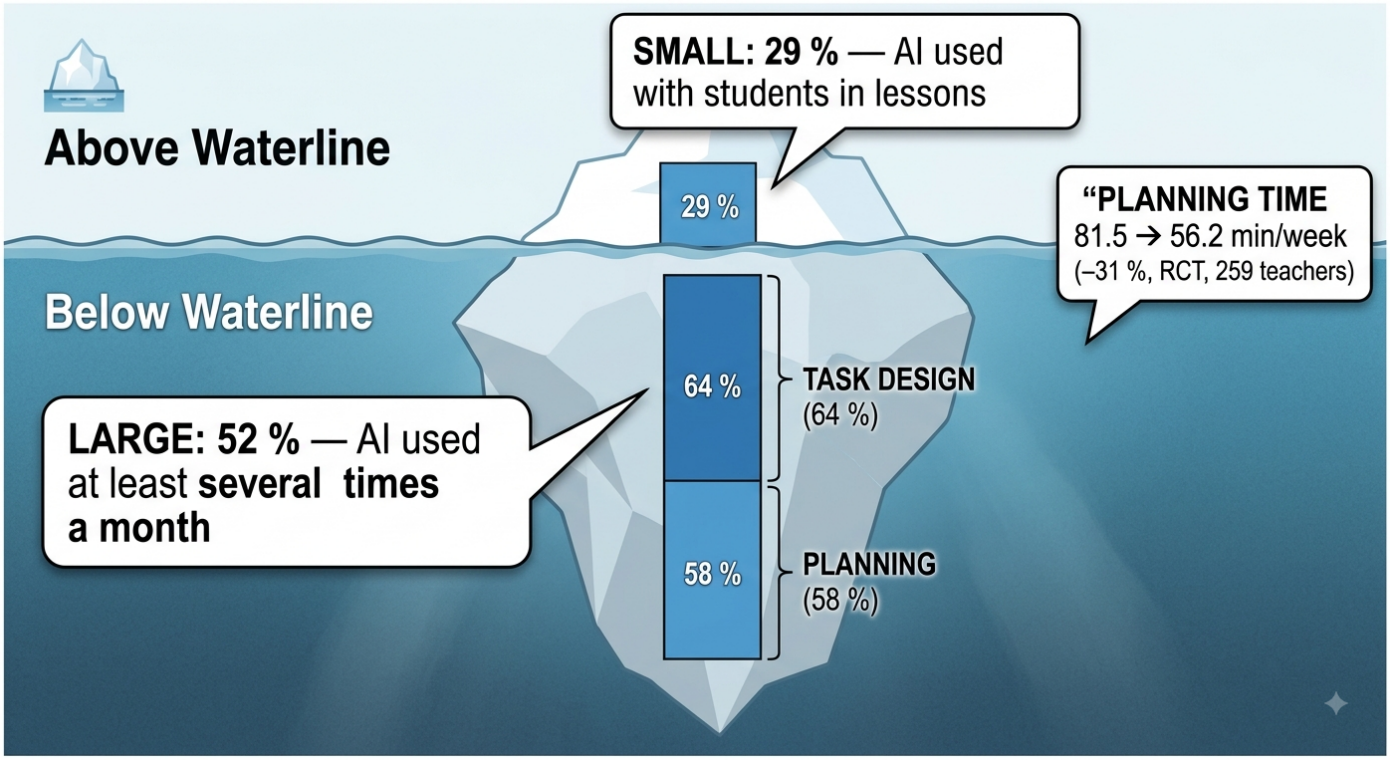

Fig. 2: Difference in AI use for preparation compared to usage in the classroom. Image created with GoogleGemini

What **is** happening is workload relief, and here the news is good. In a British trial, 259 teachers across 68 schools were given AI support for planning science lessons for 11- to 13-year-olds. Planning time fell **from 81 to 56 minutes a week**. When independent assessors reviewed the resulting materials, they found no drop in quality. Teachers' own estimates of the time they save across all subjects are even higher, at around six hours a week for regular users.

And one pattern is worth pausing on. In a large American survey, **teachers in schools with a clear written AI policy reported saving about a quarter more time than teachers in schools without one.** That is the teachers' own account rather than a measurement, and a survey cannot show which way the arrow runs. Well-organized schools may simply do both things well. But it is the opposite of what most people expect, and it is worth noting that nothing in the sources behind these points the other way: wherever this has been examined**, clear rules and heavy use go together.** -> 1.5 Compliant use of AI

Why the tool is not the point

If you wanted a ranking of the best AI tools, the research will disappoint you and then, on second reading, set you free. When researchers gather many separate studies and ask what separates the classrooms where AI helped a great deal from those where it barely helped at all, the answer is never the brand of software. It is what the students were asked to *do* with it.

There are three comparisons that really drive the point home.

- Students who used AI to transform something, e.g., rework a draft or build on their own attempt, gained about two and a half times as much as students who simply held a conversation with a chatbot.
- Students who used AI in pairs gained about 70% more than students who used it alone.
- AI was used to provide feedback on student work, and it's fair to say it made one of the biggest gains in this area. In contrast, using AI to incorporate gamification into lessons produced a gain too small to call proven..

Pedagogy beats technology, and by a wide margin. This is what happens every time somebody sorts the studies by what the students were actually doing. Two limits should be considered equally.

Good feedback is not the same as feedback that works. In one experiment, 70 university students wrote an argumentative essay and revised it a week later. They were assigned at random to three groups: one got feedback from an experienced teacher, one from a chatbot given a plain instruction, and one from the same chatbot prompted to reason through the essay step by step. Marked against the same rubric, the step-by-step chatbot produced the **best feedback of the three.** The feedback was better than the plain chatbot and better than the teacher. Then came the revisions, and **all three groups improved by much the same amount.** The best feedback in the study brought no more improvement than the lowest-rated feedback. **Better comments did not produce a better essay.**

That deserves a moment because it challenges a common assumption that better feedback leads to better learning. It does not, on its own. Feedback only does anything if it is **taken up**: read, understood, acted on, and understood well enough that the student could get it right unaided next time. **Uptake is the bottleneck, not quality**.

Two things appear to get in the way, and both are usable.

More isn't always better. In a second study, with 60 Year 11 students and 240 pieces of writing, the AI produced way more content-related comments than the teachers did. 333 against 241. It was the teachers' comments that the students acted on more often. When confronted with thirty precise corrections, a student might either stop diagnosing and start triaging or just comply. If a student follows the rules, they're basically saying that the revision was unnecessary. It's the satnav again, but on a different level. It gives you really detailed feedback, like 'Turn left in 200 meters' for your own writing.

Who says it matters. We're not really sure why teacher comments are taken up more often. This is all about interpretation, not measurement, but the most likely explanation is that feedback is seen as an act between people. What someone who knows what you did last term and will see what you do with it has to say is different from the same sentence generated on request.

What follows for practice. All the evidence we've seen on automated feedback shows that it works best when students take what they've learned in one task and apply it to a new task. It's least effective when they just copy and paste what they've learned straight from the text in front of them. And it's way more effective when you use it over a longer period of time, like a whole term,

rather than just in one go. **So, we shouldn't be asking whether the AI's feedback is good enough**. **We should be asking whether this student will ever have to use it again. If the answer is no, better comments are unlikely to help.**

And the same split shows up at scale. One team pooled 228 studies: AI clearly and consistently improved what students could produce, while any gain in their ability to plan, monitor and correct their own learning was small enough to have been chance. Programming shows it sharpest. Students with an AI assistant were measurably faster, and no better at the end than those without one. **Getting more done and getting better are not the same thing, and AI separates them.**

Europe as mirror

Across Europe the share of teachers using AI runs from **89 % in Czechia to 14 % in France.** Please consider that the surveys differ in method and timing, so read the spread, not the individual figures.

France is the instructive case. It built one of the continent's most ambitious state AI programs, an AI tutoring scheme in maths and French. Nevertheless, France still sits at the bottom of that table, as only 9 % of French teachers have received any training in AI, compared with roughly 38 % across comparable countries. **Providing tools is the easy half but qualification is the half that decides.**

Two countries have written the satnav problem into policy the way a driving school would: map first, device later. Norway recommended in June 2026 that generative AI remain essentially out of primary school and be taught only in lower secondary school, where teachers have been trained, and AI should be taught as a deliberate skill. Luxembourg stages it similarly. Note carefully: Norway issued a **recommendation**, not a binding ban, however the international press reported it.

The blind spot

One risk of AI is rarely named, and the image reaches it too: knowledge about the London city map **is not permanent.** A driver who follows the arrows for five years does not keep the city in her head. It fades, and she stops noticing the dug-up street. Your expertise is maintained by doing the very things AI is best at absorbing: writing the task, anticipating the misconception and drafting the explanation. **Where AI takes over the work through which judgment forms, relief turns into deskilling. In a German expert panel, 65 % of education experts identified this risk for the teaching profession, rising from 44 % the year before.**

Use the satnav. Just keep the map.

1.3 Responsible Use of AI

Take Home message

- **Responsibility is a design task, not a control task.** Detection of AI does not work properly.
- **And AI detection is unfair when it fails.** There is a risk of mistakenly identifying answers as AI-generated, or vice versa. Several European countries have already ruled out these tools.
- **Confident invention and inherited bias follow from how these systems are built and judged.** The consequence should be: teach them rather than waiting for a fix.
- **The working principle is: think first, then AI.** Making people commit to their own answer beats explaining the machine; building limits into the software beats banning it.
- **Prompts alone don't hold.** A chatbot tasked with acting as a tutor tends to drift back to giving answers, which makes reliable limits a purchase question.
- **Amplification is the default; equalization must be built** from access, design and scaffolding together — and the scarcest of the three is a trained teacher.{{rev}}
- **Inclusion is the clearest win**, and where teachers themselves see the strongest case.

> If you remember only one sentence from this page: **The dams are breaking. Teach swimming.**

The reading journal, second look

Back to our ninth-grade student and his suspiciously polished reading journal. The instinct is almost universal, and it is the instinct of a good teacher: *find out whether he used AI.* Build a better dam.

This chapter is about why that instinct, on its own, leads nowhere. We will come to what works instead.

Start with the uncomfortable part. **Teachers cannot reliably tell AI text from student text.** In a German study, 289 teachers of different experience were given a mix of student writing and chatbot writing and asked to sort it. Neither group could. But both groups were confident they could. Experience made almost no difference. At scale, it is worse. At one university, researchers submitted chatbot-generated answers into the real examination system through the normal channels, without notifying the markers. **94 % were never questioned**. Moreover, on average, they were graded *above* the work of actual students. So, what's the consequence? We reach for the machine! And here the finding stops being awkward and becomes an ethical problem!

The detector that fails the wrong students

Researchers tested seven of the most widely used AI detectors on TOEFL essays written by students learning English as a foreign language. Alongside, they put essays by native-speaking American eighth-graders. The native speakers' work was classified correctly almost every time. Of the foreign-language essays, **more than six in ten were wrongly flagged as machine-written** .

The mechanism is not mysterious. These tools do not detect the work of Large Language models, they detect *predictability*. Text that uses common words in expected orders scores as artificial. Anyone writing in a second language, with a smaller vocabulary and safer sentence structures, produces exactly that signal. **The tool's mistakes fall systematically on the students who are already carrying the most.** Institutions have done the arithmetic. Vanderbilt University pointed out that even if such a tool were wrong only 1 % of the time, that would still mean roughly **750 wrongly accused students a year** out of its 75,000 submissions. Consequently, they switched it off. The guidance issued by Germany's federal education ministry is blunt: the technology is "*not sufficiently reliable to provide legally secure proof.*"

Much of Europe has stopped debating this. France, the Netherlands and Switzerland advise against detectors or forbid them. The United Kingdom's exam regulator has deliberately chosen human judgment over software. The Dutch argument goes one step further and is worth carrying home: **pasting a student's work into a detector is itself an act of processing their personal data** , and doing it without permission breaks data protection law. → 1.5 Compliant Use of AI

It seems that the cheating panic is thinner than the noise around it. Researchers surveyed students at three secondary schools before ChatGPT existed and againafterward. Cheating stayed broadly where it was. Most students rejected having a chatbot simply produce their work, while thinking it fair to use one to get started or to have something explained.

The real question was never, "How *do I catch him*"*?* It is: **what does a homework task still measure, once its product can be generated on demand?**

Two things that will not be fixed, so they must be taught

Confident invention is built in. These systems generate plausible continuation of a text, which are not the same as true ones. Researchers have now formally shown where this comes from: the standard tests used to judge Large Language model reward confident guessing over admitting uncertainty, so as long as they are measured that way, a certain rate of fluent, well-formed falsehood persists. It is a property of how they are built and graded, not a bug awaiting a patch.

The slants in the training material are built in too. Moreover, they do not shrink as the technology improves. A European study of gender stereotyping across languages found that bigger

models, even after the extra training intended to make them safer and fairer, sometimes stereotype *more* rather than less. If a property cannot be engineered away, it becomes curriculum. And it can be taught: in one Finnish programme, the share of students who could identify bias in an AI system's output rose from **7 % to 44 %**. The consequence for teaching must be training error analysis instead of error avoidance. In other words: the machine's failures become the lesson.

What works: design, not control

The evidence points to one simple principle that's easy to remember: **think first, then AI.**

- **Make the human commit first.** In a series of experiments, people who had to write down their own answer *before* seeing the machine's suggestion caught far more of its mistakes than people who were shown detailed explanations of how the machine reasoned. Explaining the AI did not help to solve that issue, but forcing the judgement did. With one honest caveat: the versions that worked best were the ones participants liked least, and they helped most those who already enjoy effortful thinking. So it is a phenomenon also known as the Matthew Effect: To those who have, more will be given. They benefit the most from this approach.
- **Build the limits in.** The chatbot from Chapter 1.3, that withheld answers and asked questions instead is the same idea in software: the students who used it lost nothing on the exam, while those with the unrestricted version lost badly.
- **Make the process visible instead of hunting the product.** Drafts, working notes, a short conversation about the work. These show you what a student can do without anyone having to prove what they did.
- **Denmark is piloting the principle:** at participating schools, students may use AI while preparing for the oral English examination, and the written examination includes a handwritten, aid-free part.

Two limits are worth knowing before you rely on this. The first: the popular advice that teachers should simply instruct a chatbot to act as a Socratic tutor does not survive contact with a long conversation. At least at the moment, the system tends to drift back into supplying answers. Reliable limits have to be built into the system, which makes them a purchasing decision rather than a prompting skill. The second: making an AI explain its reasoning sounds like the obvious safeguard, but explanations do not reliably prevent over-trust. In one study they left people's acceptance of *wrong* advice unchanged, and explanations that are too technical or too simplified can increase misplaced confidence. Explanation is a condition, not a solution.

Attitude: two failure modes, not one

We worry that teachers are placing too much trust in AI. However, we should be equally concerned about the reverse. Even experienced people can be caught out by over-trusting. In one study, teachers reviewed AI-generated grades. They judged 42% of the machine's feedback to be vague or wrong, but changed only 9% of it. **Spotting an error and correcting it are different processes.**

The opposite pole is concealment. Teachers hide their own use of AI from colleagues and students, forbidding them to do what they do themselves: 77% use AI privately, while 57% say students should never use it to generate ideas. Meanwhile, a study of academically able students in grades 6 to 8 found that between a third and two-fifths already said the AI knows more than their teachers. **A profession that hides its practices cannot model responsible practice**, and transparency is not just a courtesy. It is the mechanism by which teaching by example works.

Does AI level the field or tilt it?

For work-related tasks, AI evens out differences. In one experiment, participants with higher levels of education performed significantly better than those with lower levels of education. When each person was given an AI assistant, these differences disappeared almost entirely. Then take the assistant away again, and much of it came back. **What the AI closed was the gap in output, not the gap in ability.**

In learning tasks, it does the opposite. One study found no average benefit at all, but a widening distance between students who already knew a lot and students who did not. Another looked at students writing in a foreign language: the weaker writers received *more* corrections from the AI and successfully acted on *fewer* of them, and several found the flood of feedback dispiriting. **Feedback without the ability to sort it does not help; it buries.**

The gap also exists before anyone touches a tool. In Germany, **80 % of students from the most advantaged homes see AI as an opportunity, against 55 % from the least advantaged**. In Britain, the divide runs through the staffroom rather than the device cupboard: **45 % of teachers at private schools have had formal training in using AI, against 21 % at state schools.** Within the state sector, wealthier schools again outpace poorer ones. A survey of children across 20 European countries finds the same social split in how much and how variedly they use AI. One line captures the whole problem better than any figure: "*the rich have access to technology and people to help them use it, while the poor have access to technology only.*" Which is the strongest available argument that teachers matter to close this gap.

So the bottleneck is what learner has a teacher who has been trained. **AI compensates when three things hold together: secure access, a tool designed for learning rather than answering, and a teacher scaffolding its use.** Remove any one, and it amplifies instead. Since all three tend to be present in the same schools and absent in the same schools, **amplification is the default and equalisation is something you have to build.**

1.4 Compliant Use of AI

Take Home message

- ***There is no legal vacuum.*** Data protection law has applied all along.
- ***The timetable moved.*** The high-risk duties for education AI start on **2 December 2027**, not August 2026. Check your school's handouts.
- ***Already binding since February 2025.*** Emotion recognition in schools is prohibited, and schools deploying AI must ensure staff competence. Labelling of AI-generated content starts in August 2026.
- ***The hardest layer is data protection.*** No personal data about pupils without a contract that excludes training on your inputs; consent usually does not carry; deleting names is not anonymizing.
- ***In copyright, the upload is the problem.*** The teaching exception does not cover feeding protected works into an AI.
- ***Detectors do not hold up as evidence.*** Reasoned inference from concrete signs does. Rules must be clear in advance, and labeling beats banning.
- ***Grades stay human*** — and rubber-stamping an AI's suggestion is not a human decision.
- ***Responsibility sits with procurement, not with you*** — *and* ***compliance enables.*** The schools with clear rules are the ones whose teachers report saving the most time.

> If you remember only one sentence from this page: **Compliance is a house, not a wall.**

The reading journal, third look

Our ninth-grade student, his reading journal, his teacher's suspicion. In Chapter 1.2, it was a question about learning; in Chapter 1.3, about professional judgment. Here it is a court file.

In December 2025 the Administrative Court of Hamburg, Germany, ruled on an urgent application in the case. The teacher had noticed that the journal did not match the boy's writing style in a class test. As a consequence, the school treated it as deception. This was followed by an urgent application to stop this, but the court refused. Undisclosed AI use in schoolwork counts as deception even where the school has no explicit rule against AI. All students must assume they are to work independently unless they have been told which aids are allowed. One caveat to keep in mind from the outset: this is a single court, and the decision is not yet final.

This is a good moment to address the most common misconception in this field. **There is no legal vacuum.** There never has been.

The house has three storeys

Most teachers envisage AI law as a single wall under construction somewhere in Brussels. However, it is more useful to picture a house that you already live in. **The foundation is data protection.** It is load-bearing. It was laid years ago. It governs your Tuesday morning. **The upper storey is the EU AI Act,** Europe's AI law, which categorizes systems according to their potential for harm and imposes stricter obligations the higher the risk. It is partly occupied and still under construction, with the completion date just pushed back. **The doors and windows are copyright and examination law** — small and specific, and touched daily without noticing.

Ground floor: data protection, clearer than many think

Handling someone's personal data always requires legal permission. As a public body, a school's permission essentially comes from its statutory duties under school law. There is no general license that comes with being a teacher.

Three consequences follow, forming the practical core of this chapter.

Consent does not provide a solution. One obvious solution would be to simply ask the parents. However, this approach has two drawbacks. Firstly, consent must be freely given. In a school context where grades are a factor, a teacher's request does not constitute a free choice. Secondly, consent must be informed, which is difficult to achieve when nobody knows where the data will end up. **Permission must come from school law**, not a signature, as this shifts responsibility upwards, away from you.

Removing names does not anonymize data. In 2024, Germany's data protection authorities jointly stated that if someone could still be identified from the context, the data is still personal. This is particularly problematic in schools because the types of text that teachers most want help with — such as essays, free writing and learning journals — identify their author by style.

The dividing line is the contract: a provider may handle your data if it does so strictly on your instructions and there is a written agreement to this effect. However, the moment it uses what you type for its own purposes, the arrangement collapses because it is no longer working on your behalf. So the practical test is this: I**s there such an agreement, and does it exclude training on your inputs?** If not, no personal data should be entered.

Two footnotes worth noting: Data relating to health, disability, and special educational needs is subject to even stricter protection under EU law, and at least one German state (Baden-Württemberg) has explicitly prohibited its use in AI systems by ministerial decree. Where private devices are involved, **the legal responsibility lies with the school, not you personally**, which relieves you and obliges your head teacher.

Upper floor: what the AI Act does and does not require yet

Here is the correction that makes most existing guidance out of date. In July 2026, the EU postponed the heaviest set of AI Act duties for education, which are those attaching to systems

classed as **high-risk**, meaning systems used to decide admissions, to assess what pupils have learned, or to monitor them during examinations, **from August 2026 to 2 December 2027**. A great many handouts written in 2025 and early 2026 still print the old date. If a document in your school does, it needs a footnote.

The more useful correction runs the other way: **what already applies is underestimated, and what is endlessly discussed does not apply yet.**

Two things have been binding since **2 February 2025**.

- **Emotion recognition in schools is prohibited outright.** Systems that claim to read pupils' feelings — attention detection, mood analytics, "engagement" scoring from a webcam — are not a grey area. They sit on the Act's short list of practices that are simply banned.
- **Schools that deploy AI must make sure their staff know what they are doing with it.** No fine is attached to this one, which is why it is widely overlooked. It is binding all the same.

Read that second duty alongside Chapter 1.3 and something clicks. The gap between knowing about AI and being able to teach with it is no longer only a professional problem. **Since February 2025 it has been a legal obligation.**

Since **2 August 2026**, AI-generated content has to be labeled as such. From **2 December 2027**, systems used to decide who gets admitted, to assess what pupils have learned, or to monitor them during examinations count as high-risk. Whoever *uses* such a system then assumes the duties that come with it. This incorporates a requirement that the people overseeing it be, in the Act's words, "*competent, trained and authorised*". Note the wording. The law has arrived at the same conclusion as the pedagogy: **oversight without expertise is not oversight.**

Doors and windows: copyright and examinations

Counter-intuitively, **what comes out is the easy part.** A text or image generated by an AI generally has no human author in the sense copyright requires, so it is usually free of copyright itself. It is not, however, guaranteed to be clean: in the first major European ruling of its kind, a Munich court found that a chatbot had memorized song lyrics and reproduced them almost word-for-word. The court placed the liability on the **provider**, not on the person who typed the request. Whether a teacher who passes on such material could be liable is, as things stand, **an open question**. What goes in is the clearer problem, but note that this is a question of **copyright, not of the AI Act.** The AI Act says nothing about what teachers may feed into a system; that is an older and entirely separate body of law.

The exception that lets teachers copy material for their classes covers exactly that: copying a limited share of a work, for that class. **Uploading a protected work into a commercial AI system is a different act, and the exception does not address it.** Material intended for school teaching, which is mainly textbooks, is excluded from the exception in any case. So scanning a chapter to have it summarised is an everyday act that the teaching exception does not

obviously cover.

How much should that worry you? Less than the gap suggests, and more than nothing. The question is **genuinely unsettled**: no higher court has ruled on it, and the specialist literature regards the existing exceptions as a poor fit for the situation. Where a rights holder objects, the claim would be a civil one, and in practice, this is a matter for school policy and procurement rather than something an individual teacher can settle. What is *not* unsettled is the narrow case, and it is the line worth remembering: **school textbook and publisher material should not go into AI tools at all — neither scanned nor retyped, not even to generate a worksheet.**

Case law in this field seems currently inconsistent across Europe. In February 2026, the Administrative Court of Kassel, Germany, decided two university cases the same way, and sharply: where a student has signed a declaration that the work is their own, the line is crossed "*already with a single undisclosed use of generative AI*". A Paris administrative court, in February 2026, forbade a university to sanction a student at all: it had produced no rule governing AI use, and without a rule set out in advance there was no disciplinary breach to find. A Swedish appeal court had reached the same result earlier, on facts almost identical to a Dutch case that went the other way. In this case, a student submitted invented sources, but the course rules permitted "all aids" in take-home work, so there was nothing to have broken. Where fabrication *was* clearly proven, and a rule existed, courts have upheld sanctions. The Dutch Supreme Administrative Court did so in 2025.

So the pattern across jurisdictions is not about countries. It is about **whether a rule existed beforehand, and whether the deception was actually shown.**

Three rules follow for practice.

1. **Detectors do not carry the burden of proof.** What the law does allow is an inference from typical signs. This includes repeated, oddly polished phrasing, a mismatch between what they can write and what they can say, and sources that turn out not to exist. The student can then explain. That is a legitimate route, and it does not require software.

   Detector output, at best, is one sign among others. The failure rate reported in Chapter 1.3 renders the instrument unusable as evidence.
2. **Rules must be clear before the assessment, not after.** Neither a blanket ban nor a blanket permission holds up. What works is a declaration of independent work with a general clause about AI, plus a duty to state what was used. In Germany, Bavaria went furthest in 2026: its universities may **not** forbid AI in unsupervised assessments at all, but must require students to declare it. Labeling instead of banning.
3. **Grades stay human.** No decision with legal consequences for a person may be made by a machine alone, and a grade is such a decision. Germany's Standing Conference of the Education Ministers puts it even more strongly: assessment is a task only teachers can perform. Crucially, **formally waving through an AI's suggestion does not count as a human decision.** So the finding from Chapter 1.4, wrong AI feedback is also a pedagogical and a legal problem.

Law, ethics and didactics therefore converge on the same answer: **assess the process, not just the product.**

The real address of responsibility

In response, Germany has introduced state-provided school AI environments. These are vetted platforms with negotiated contracts, rather than leaving teachers to sign up to whatever is free. While these solve problems that individual teachers cannot solve alone, they are still criticized for how they are governed, whether teachers actually accept them and whether the assessments they support are sound. Since the rollout is currently limited to a select group of schools and specific subjects, we will need to monitor over the next few months how teachers are faring with the platforms.

What is distinctive is narrower than it may initially appear, and it is worth stating precisely: Rather than putting it out to tender or buying access to someone else's, Germany is the only European system where a state body builds and runs the platform itself. France also provides a platform for an entire year group, but it commissioned one and two private companies deliver it: one wrote the software and the other hosts the data. Most of the rest of the continent has simply purchased one. Slovakia has gone the furthest, with 20,000 licences for teacher training faculties that will be extended towards some 80,000 school teachers by 2026. The state will only pay for the licences actually used, and the provider is contractually barred from training on Slovak school data. Estonia negotiated a similar arrangement for its upper secondary schools, and Northern Ireland invested £10.7 million in licences for all teachers in all schools. The contractual guarantee that your inputs are not used for training — the feature that the German platform was designed to provide — can be purchased.

Building it yourself does not automatically give you more control. The German platform runs on the same commercial language models as everyone else's on European servers, as an industry association has pointed out. By contrast, the French system is hosted in France under the highest security certification issued by the French state. One country built the surface and rents the engine, while the other bought the software and keeps the keys to the building. Neither option is clearly the most sovereign choice, which is worth bearing in mind before using the term.

The Netherlands took a different approach, and it is worth knowing about because it offers a more honest answer to the same question. Rather than building or buying, Dutch schools and universities joined forces and had the major providers formally assessed. When they examined Microsoft's AI assistant, they found two issues that could not be resolved: the criteria for an automatic content filter that the provider would not disclose and the retention of eighteen months of diagnostic data that could not be justified. They negotiated. In March 2026, however, the provider declined to change. The assessment therefore remains amber, and Dutch schools are advised to exercise caution and use their own judgement on a case-by-case basis rather than granting clearance. One country removed the risk by building around it, while another took a closer look and decided it was not ready yet. Both approaches offer solutions to the problem set out at the start of this chapter.

The comparison also settles one more thing. The complaint that teachers prefer the free chatbot to the official one is not due to a failure of implementation in Germany. In Estonia, where the state negotiated access for every upper-secondary student, barely more than a third were using the official chatbot weekly after a year. In turned out, that most of them had been using free tools before the program started. **Everywhere, the vetted environment competes with the more convenient one.** Providing the official tool is necessary. But this is not sufficient.

However, the structural point stands: compliance is an enabling condition, not a hindrance.

Remember what was said in Chapter 1.2? **The schools with clear rules were the ones whose teachers said they saved the most time**. There is nothing in the sources behind these chapters that suggests otherwise.

1.5 Current State of AI in Education

Take Home message

- **AI is already widely used in educational contexts.**
  - **Teaching with AI.** AI is currently used by teachers to prepare or improve lesson plans, for assessment and to distinguish content across different educational levels.
  - **Teaching about AI.** AI is used by students though they have no groundings on what is AI, which are its limitations and to use it responsibly.
- **Technical understanding and critical thinking are needed.** Teachers and students need to understand how language models work (training, data handling, predictive nature) and develop practical verification skills, from spotting fabricated citations to fact-checking AI-generated content.
- **Responsible use and ethical stance are to be defined**. A grasp of the legal and ethical landscape (AI Act, GDPR) is needed, along with an approach to AI as a support rather than a substitute for human thinking, grounded in transparency and accountability, plus explicit instruction in prompting as a teachable skill.
- **Equity and new assessment models are needed.** Equitable access to AI tools needs safeguarding (public funding, open models, minimum age policies), alongside a shift in assessment from grading the final product to evaluating the process and students' ability to critically engage with AI output.

> If you remember only one sentence from this page: **AI is already present in educational contexts but is not regulated. Structured frameworks are needed to support educators' professional development improving their knowledge and letting them teach with AI and about AI.**

Current Uses of AI in Educational Contexts

Nowadays, artificial intelligence (AI) is embedded in everyday educational practice, well beyond pilot projects. Common applications include adaptive learning platforms, intelligent tutoring systems, automated grading, and AI-assisted lesson planning for teachers (OECD, 2026). The most visible shift, however, is the direct use of general-purpose tools by students and teachers alike.

Teaching With AI

AI is increasingly used by teachers as a support tool across multiple stages of their work, from lesson planning to grading. According to OECD's TALIS survey, 37% of lower secondary teachers reported using AI in 2024, with 57% saying it helps them write or improve lesson plans (OECD, 2026). Beyond lesson design, teachers are using AI to differentiate content for different skill levels, generate practice materials, and support the grading process. At the same time, adoption is not without friction: 72% of teachers expressed concern that AI could enable students to pass off AI-generated work as their own, reflecting a broader tension between AI's practical benefits for teaching and its implications for academic integrity (OECD, 2026).

Teaching About AI

Alongside using AI as a teaching aid, there is a growing need to teach students about AI itself: how it works, its limitations, and how to use it responsibly. This includes building a basic understanding of language models as prediction engines rather than databases, which explains their tendency to hallucinate, as well as practical skills such as fact-checking AI output and recognising fabricated citations (UNESCO, 2024). It also involves developing critical thinking around prompting, shifting students from simply following instructions to formulating them, and instilling awareness of the ethical and legal dimensions of AI use, including data privacy and the EU AI Act. UNESCO's *AI Competency Framework for Teachers* (2024) frames this as a core professional competency, arguing that teachers themselves need training in these areas before they can effectively teach them to students.

What is missing?

Alongside institutional frameworks and adoption data, it is important to understand how these dynamics play out in practice. A series of workshops was conducted in the context of the DUAL.AI.TEACHer project. Numerous education stakeholders were called to gather first-hand perspectives on the current use of AI in teaching. These workshops surfaced a set of concrete needs and concerns expressed by educators, covering practical, pedagogical, and training-related aspects. The following sections outline these needs in detail.

- **Understanding How AI Works.** Both students and teachers need to understand how language models are built and trained, and what happens to their data once provided. They also need to grasp the nature of these models as prediction engines rather than databases, which explains their tendency to hallucinate. Teachers need a simple high-level course covering the basics of how LLMs work.
- **Critical Thinking, Questioning and Verifying AI.** Students and teachers need to be able to question AI-generated answers and to recognize fabricated information. Practical verification skills are also needed, from checking AI output against official sources to spotting AI-generated images and videos. Students need to shift from following instructions to formulating them, thinking critically about the desired output.
- **Ethics, Data Protection and Regulation.** A general understanding of the ethical and legal aspects of AI use is needed, including the EU AI Act and GDPR. Safe and responsible

use is essential to prevent illegal practices along with awareness of related issues like bias in teaching materials and the unreliability of AI detectors.

- **Responsible Use.** AI needs to be used mindfully as a support, not a replacement, leaving room for students' own thinking. Transparency matters, with students expected to flag which parts are their own, echoing "cyborg writing", where AI extends rather than replaces the writer's thinking. Student ownership and accountability are also needed.
- **Practical Skills.** Prompting needs to be treated as a teachable skill, not something acquired independently. Both teachers and students need explicit instruction in writing effective prompts, and teachers need training in prompt engineering for lesson planning, differentiation, and grading. Broader tool familiarity is also needed to match tools to subjects.
- **Equity and Infrastructure.** Access to AI is at the centre of growing equity concerns, as wealthier districts can afford better tools, widening competence gaps between schools. Public funding is needed to guarantee equal access regardless of a school's resources, preferring open AI models over corporate ones as part of this effort. A minimum age for AI use also needs to be defined.
- **Assessment.** A different approach to assessment is needed, focused on process rather than final product, including students' reflections on AI-generated drafts. AI-integrated assignments are also needed, grading students on their ability to critique, fact-check and edit AI output, raising questions about the human role when teachers themselves rely on AI for grading. Concrete classroom methods are needed too, such as fact-checking AI text, comparing chatbot responses, debating AI use in homework, and incorporating non-digital tools like flip charts.

1.6 Legislation

Take Home message

- **The EU AI Act is the legal basis.** The EU's rulebook for AI is in force since August 2024 and phasing in through 2027. It applies not just tech companies, but also to schools.
- **AI literacy is a duty, not a course to tick off.** Schools must support staff in building AI literacy.
- **Several AI uses are banned outright.** Among them: inferring students' emotions from facial or vocal data, AI that exploits children's vulnerabilities to distort their behaviour, and building facial-recognition databases or inferring traits like race or religion from biometric data.
- **Risk depends on the decision, not the product.** Ordinary writing, research, translation and lesson-planning tools are not high-risk simply because they are used at school. They become high-risk once they influence admission, learning pathways, exam outcomes or employment decisions.
- **Not everything AI-generated needs a label.** The disclosure duty mainly sits with the provider; teacher-reviewed material with human editorial control is exempt.
- **High-risk AI needs a human who can overrule it.** AI output is never the final decision. You must understand its limits and be able to disregard or override it.
- **The AI Act sits on top of other law, it doesn't replace it.** National data protection, education, employment, child protection and copyright law continue to apply alongside it.
- **The GDPR governs the data, not just the tool.** Even an AI system that is fine under the AI Act can still be off-limits if it processes personal data about students or staff without a valid legal basis.

> If you remember only one sentence from this page: **Use AI to support your professional judgement, not replace it—stricter rules apply as soon as AI helps make decisions about grades, educational pathways or people.**

The EU AI Act

The EU AI Act (Regulation (EU) 2024/1689) is the EU's single, risk-based rulebook for artificial intelligence. Instead of regulating specific products, it classifies AI systems by the risk they pose and attaches obligations accordingly — from an outright ban for a small set of practices, through extra duties for "high-risk" uses, to no special rules at all for everyday tools. It applies directly in every member state, to public and private organisations alike.

The Act phases in step by step. It has applied since 1 August 2024; since 2 February 2025, the banned practices below and the duty to build AI literacy (Art. 4) are already binding; from 2 August 2026, most of its remaining provisions apply.

Who is responsible for the use of AI?

If a school introduces and controls an AI system, the school or school authority is usually the *deployer* (Art. 3(4)). The teacher is normally not a separate deployer when using the tool under the school's instructions, but must still follow those instructions and all relevant school and data-protection rules.

Teachers need to know what they are using

Schools and school authorities must take appropriate measures to support the development of AI literacy among those who use AI systems on their behalf (Art. 4). The AI Act does not specify a particular level of competence, certificate or training format. These rather depend on the AI system used, how it is used, and who may be affected.

AI practices banned outright for schools

The AI Act bans eight AI practices outright (Art. 5); four of them are directly relevant to schools:

1. Using AI systems in educational institutions to **infer emotions or intentions from biometric data** (such as students' facial or vocal characteristics) is generally forbidden. Narrow exceptions exist for medical or safety purposes (Art. 3(39); Art. 5(1)(f)).
2. AI systems that exploit someone's vulnerabilities due to age (as with children), disability, or social or economic situation, in a way that distorts their behaviour and causes them harm, are banned regardless of intent (Art. 5(1)(b)).
3. Building or expanding a facial-recognition database through untargeted scraping of images from the internet or from CCTV footage is banned. This is relevant if a school considers a facial-recognition tool for attendance or exam monitoring (Art. 5(1)(e)).
4. Using biometric data to categorise people in order to infer protected traits such as race, political opinions, religion or sexual orientation is banned. This is relevant for any biometric-based tool, such as exam-proctoring software, that goes beyond its stated purpose (Art. 5(1)(g)).

Does AI-generated content always need a label?

If students interact directly with an AI system, such as a chatbot, they must be informed that they are interacting with AI. Usually the system provider is responsible for making this clear (Art. 50(1)).

Schools and teachers must disclose deepfakes, or AI-generated text published on matters of public interest, when these are used in the classroom (Art. 50(4)). This requirement does not apply where the AI-generated content has undergone effective human review or editorial control, and a person or institution assumes editorial responsibility for its publication. An AI-assisted worksheet reviewed by a teacher therefore does not automatically require a label.

When AI in education becomes high-risk

Ordinary writing, translation, research and lesson-planning tools are not high-risk simply because they are used for work. Classification depends on the system's intended purpose, not simply on which product is used (Art. 6 · Annex III).

AI systems may be classified as high-risk when they are intended to:

For Teachers	For Schools
• decide about students' access, admission or assignment to educational or vocational institutions and programmes; • evaluate students' learning outcomes or steer their learning processes; • assess the appropriate level of education a student will receive or can access; or • detect students' prohibited behaviour during examinations. *(Annex III, point 3)*	• recruit or select teachers, including screening applications; • decide on contract terms, promotion or termination of a teacher's employment; • allocate tasks based on a teacher's behaviour or personal traits; or • evaluate teachers' performance. *(Annex III, point 4)*

A limited checking or preparatory tool may also fall outside the high-risk category if it does not significantly influence a decision. However, an education system that profiles individuals is always considered high-risk.

What high-risk AI means for schools and teachers

Classification of an AI system as high-risk does not automatically rule it out from use in schools. But high-risk systems come with obligations (Art. 26):

These obligations for high-risk AI in education apply from **2 December 2027** — but the prohibitions above and the AI-literacy duty already apply today.

For Teachers	For Schools
• use it only for its approved purpose and follow its instructions; • do not treat its output as a final decision; • understand its main limitations and likely sources of error; • watch for overreliance on apparently convincing results; and • be prepared to disregard or override its output.	• provide competent and authorised human oversight; • ensure that any input data under its control is relevant and sufficiently representative; • monitor the system and respond to risks or serious incidents; • keep automatically generated logs for at least six months where those logs are under its control; and • inform students, teachers or affected people in advance when the system assists decision-making about them.

Public bodies and private organisations providing public services may also have to complete a fundamental rights impact assessment before using high-risk AI (Art. 27). Public authorities and organisations acting on their behalf may additionally have to register the system in the EU database (Art. 49). Whether this is the responsibility of the school or its governing authority depends on the national school system.

The AI Act is not the only law that matters

The AI Act does not replace european or national regulations on data protection, education, employment and worker participation, child protection or copyright (Art. 2(7)). For example, the fact that a tool is not high-risk does not mean that you may upload students' work or personal data to it.

A tool may be acceptable under the AI Act but still prohibited by data-protection law or your school's rules.

The GDPR

The GDPR (Regulation (EU) 2016/679) is the EU's law on data protection. It plays a supporting role alongside the AI Act: while the AI Act asks whether and how an AI system may be used at all, the GDPR asks whether this particular data about this particular person may be processed this way?

- **It applies as soon as personal data is involved.** Names, school email addresses, student work, voice recordings, transcripts, photos, grades, attendance and learning-progress data all count as personal data (Art. 4).
- **Some data needs extra protection.** Health information, biometric data and data revealing racial or ethnic origin, religion or similar traits are "special category data" and are, in principle, off-limits without a specific exception (Art. 9) — one reason the AI Act

separately bans several biometric-inference practices in schools.

- **Know who is responsible.** The school is normally the “controller”; an AI provider processing data on the school’s behalf is a “processor” and needs a data-processing agreement that spells out its obligations (Art. 28). Check this before rolling a new tool out school-wide.
- **Data leaving the EU needs a safeguard.** Many popular AI tools are hosted outside the EU. Personal data may only travel there if an adequacy decision or another approved safeguard, such as standard contractual clauses, is in place (Art. 44).
- **Students and staff keep their rights.** They can ask what is held about them, have it corrected or deleted, and object to decisions based solely on automated processing that significantly affects them (Art. 15–Art. 22) — a right that runs alongside, and reinforces, the AI Act’s human-oversight requirements for high-risk systems.

Where to find help

Enforcement is organised differently in each EU country. Because national responsibilities may change, use the European Commission's current list of national market-surveillance authorities. Anyone (a teacher, a student, a parent) can lodge a complaint with that authority if they believe the Act has been breached (Art. 85). Where a high-risk system, such as an admission or exam-monitoring tool, has led to a decision that significantly affects a student, they (or their parents) can also ask for an explanation of that individual decision (Art. 86).

National rules for your country

The AI Act is EU-wide, but enforcement bodies, school law and data-protection practice differ by country. Country-specific notes for the Playbook's partner countries will be added here:

- Austria
- Czechia
- Germany
- Latvia
- Poland
- Slovakia
- Slovenia
- Spain