Department for Education

The safe and effective use of AI in education

Module 3 – Developing the safe use of AI in education video transcripts

June 2025

Contents

Video 1 – Module introduction

Presenter: Welcome to module three, video one from the Safe and Effective Use of AI in Education online resources.

In this module, we'll explore the risks and safeguarding considerations for educators related to generative AI. We will also look at what we can do to mitigate them. There will then be a short task for you to apply the knowledge you've gained from this module, followed by further videos examining some of the risks in more detail.

There will be multiple-choice questions so you can test your knowledge and take a moment to reflect.

We often find that children and young people adopt new technologies ahead of adults. Generative AI is no exception. According to Ofcom's 2024 Online Nation study, 54% of online children in Britain aged 8 to 15 had used a generative AI tool in the past year, rising to 66% of teens aged 13 to 15, compared to only 41% of UK internet users aged 16 plus.

The most commonly used generative AI tool among online children was ChatGPT. This was used by 37%, followed by Snapchat My AI.

As with all technologies, this brings some challenges. When it comes to academic work, there's no doubt that some students and pupils are using AI tools. Over half of 8 to 15-year-olds in Britain who used generative AI said they would use it to help with schoolwork, according to the Ofcom Online Nation report.

This suggests we must be very clear about our expectations of generative AI use with students and pupils. We'll dig more deeply into that in the last video in this module.

There are many positive use cases for generative AI systems, but as educators, we must understand the risks. Here, we'll explore five key areas:

- Pupil or student exposure to inappropriate or harmful content
- Pupil or student exposure to inaccurate, misleading, or biased content
- Data protection breaches
- Intellectual property infringements
- Academic integrity challenges

Video 2 – Key risks associated with generative AI

Presenter: Welcome to module three, video two from the Safe and Effective Use of AI in Education online resources.

In the first video of this module, we had an introduction to some of the key risks related to AI and education, and in this video we'll explore some of the risks in more detail.

It's important to remember that if you encounter any incident that is a safeguarding concern, whether or not it's related to AI, you must follow the statutory guidance Keeping Children Safe in Education, which includes protecting children online, as well as your organisation's safeguarding procedures. The advice in this toolkit is not meant to replace your safeguarding training but explores some of the risks that are specific to generative AI systems.

In the DfE's generative AI product safety expectations, the Department for Education provides clear guidance that outlines the capabilities and features that generative AI products and systems should meet to be considered safe for users in educational settings. Key areas of the guidance include content filtering, activity monitoring, security, data protection, intellectual property, transparency and accountability. We will cover some of these in this module in more detail.

When considering online safety, we know that the online world often mirrors or even amplifies the same behaviour issues we see offline. For instance, online bullying can be much more pervasive and difficult to stop than offline bullying. With this in mind, it's important for school leaders to regularly update safeguarding policies and behaviour policies to reflect the rapidly changing risks from AI use. There is further guidance on this in the leadership toolkit.

The UK's Online Safety Act puts a range of new duties on social media companies and search services, making them more responsible for their users' safety on their platform.

It's important to remember that safeguarding is everyone's responsibility. As a teacher, lecturer, or other member of school or college staff, it is essential to understand your role in keeping children safe online. Keeping Children Safe in Education highlights the importance of online safety training for all staff, ensuring you are aware of the risks and your responsibilities. You should familiarise yourself with your school's filtering and monitoring systems and know what to do if you have concerns about a child's online activity. Your designated safeguarding lead or DSL should be your go-to for guidance on online risks, and regular training will help you to stay updated on emerging threats.

Being informed about online safety in your teaching and conversations with students will help to increase awareness of online dangers. By taking a proactive approach, you contribute to a whole-school safeguarding culture that protects students both in and out of the classroom.

The content we create using generative AI tools can be very useful. However, as with all systems, there is the potential for misuse. In module two, we learnt that AI systems are trained on large amounts of existing data, and so they may reproduce bias that already

exists in society, or replicate bias introduced by the people who designed the AI model. This means we need to critically evaluate the output to make sure it's appropriate for our use.

We've previously explored some of the unintentional inaccuracies known as hallucinations that result from the way generative AI system works. With AI audio, image and video generation tools, people can generate content that is misleading or harmful. You can make it appear that someone said or did something that they didn't, and this is known as a deep fake.

The rise of AI generated child sexual abuse images presents a serious safeguarding threat, with predators using AI to manipulate real images of children, create deep fake content, and generate disturbingly realistic abuse material. These images not only exploit children, but also revictimise survivors of past abuse. Offenders are using AI to groom, blackmail and coerce children, making detection more difficult and increasing the risks our pupils and students face online.

In response, the UK is introducing new laws that criminalise the possession, creation and distribution of AI tools designed to generate child sexual abuse materials. The government, alongside charities and safeguarding organisations, stress that AI misuse must be tackled alongside its benefits, ensuring technology does not outpace child protection. Teachers play a key role in safeguarding by staying informed, recognising risks and supporting pupils in navigating the online world safely.

The fact that digital tools required to generate these images are freely available makes it more difficult to detect the creation of these materials. It also increases the potential for young people to be able to generate images they might otherwise be blocked from accessing online. Schools need to ensure that filtering and monitoring systems are effective in dealing with these risks.

In addition, it's possible for generative AI technology to be used to create very realistic avatars and chatbots that can simulate human conversation and interaction. This can make it almost impossible to distinguish between a real person and an avatar, and this has the potential to be used for grooming.

The UK's 2023 counterterrorism strategy highlights that emerging technologies, including generative AI, present both risks and opportunities for counterterrorism efforts. While the full impact of generative AI on terrorist activity online is still unfolding, there is already evidence of early experimentation, such as using AI to generate synthetic propaganda, also known as deep fakes. Although we have not yet seen widespread adoption, we know that these technologies are going to become more sophisticated, and so there's a risk that extremist content will become even more convincing and harder to detect. If combined with other tools, AI-generated propaganda could be produced and distributed at an industrial scale, potentially overwhelming content moderation systems and increasing the challenge of countering online radicalisation.

To address these concerns, the UK is actively working to understand the risks associated with generative AI and exploring mitigation strategies. This includes monitoring how extremists may try to exploit AI for recruitment, misinformation and incitement, while also

considering how AI can be used as a tool to strengthen counterterrorism efforts. Educators have a vital role to play in supporting young people to develop critical thinking skills, enabling them to recognise AI-generated misinformation, question extremist narratives and navigate the online space safely. This aligns with the Prevent duty, which requires schools and colleges to safeguard students from radicalisation by building their resilience to extremist content and online influence.

Lauren Thorpe, Chief Transformation Officer, United Learning:

“We've got to think really carefully about how we safeguard people, both in terms of protecting their data – and that's not just their sort of personal data but also protecting what they produce – because we know that when those things go into an AI model, they sort of become lost to you in terms of your ownership of that information. So we need to think really carefully about that. We also need to think about the impact that AI outside of school is having on us as human beings, so the way that it's being used in social media, the way that we're interacting with other things and it's influencing what we think because the AI is showing us more things that it thinks we like.”

Presenter: To help to mitigate this, schools and colleges also need to ensure that their online safety policies are up to date and make sure we always follow safeguarding procedures as set out in Keeping Children Safe in Education. As educators, we must all be aware of the potential for AI tools to be used to generate misleading or harmful content, and look for opportunities to enhance pupils' and students' understanding about misinformation and deep fakes alongside the laws that are in place.

Critical thinking and AI literacy are the terms often used to describe approaches that can be promoted within the existing curriculum. In addition, the National Centre for Computing Education has produced curriculum resources that support an understanding of online safety, AI and the associated risks.

You saw in earlier modules that some AI models are trained with texts that is inputted as a prompt, and this means we need to be particularly careful about the information we enter as prompts. Entering sensitive or personal information may result in a data breach, which should be reported promptly to your data protection officer. You can mitigate this by only using AI tools for work that are provided by your school or setting and which have appropriate safeguards in place.

The Department for Education's data protection in schools project developed a video on offering guidance on protecting children's privacy when using AI. You should avoid entering any personal information or student work into an AI system unless your setting has said it's safe to do so and appropriate permissions are in place. It's also very important to be transparent about how you're using the tools.

When using AI, it's crucial to understand the difference between data protection and intellectual property – two distinct legal considerations that sometimes get confused.

Under UK GDPR, if you're processing personal data – that's information that identifies an individual, such as pupil names, assessment or work, for example – you must have a lawful basis. This could be consent, contract, legal obligation, vital interests, public task or legitimate interest. Before using any AI tool that processes personal data, you must

ensure that it complies with UK GDPR. If the tool stores, learns from or shares the data, you could be breaching data protection law unless appropriate safeguards are in place.

Intellectual property or IP is a separate issue. IP law protects original creative work, including lesson plans, teaching resources and pupils' and students' work. If you've used an AI tool that trains on uploaded content, you need to ask, do I have the legal right or necessary permissions to use this material and who owns the copyright?

For instance, under UK copyright law, a pupil or student automatically owns the copyright of their written work, artwork or other creative outputs. If you upload their work into an AI system that stores or reuses it, you may need their explicit permission or the permission of a parent or guardian, depending on the circumstances.

To summarise, data protection is about safeguarding personal data and ensuring that you have a lawful basis for processing it. Intellectual property law governs ownership and permission for using copyrighted materials, including learning resources and pupils' or students' work. Before using AI always check both to ensure compliance.

In the case of student or pupil work, we could be infringing intellectual property rights if, for example, we were to enter their work into an AI system for marking or feedback. Intellectual property covers a broader range of creative and original works, including text, images, music and code. While personal data laws focus on protecting individuals, IP laws protect the rights of content creators and copyright holders.

Before using AI to generate or manipulate content based on someone else's work, such as using an LLM or large language model for marking or feedback on students' work, permissions must be obtained from the rights holder. This distinction is crucial because an individual's work may be protected even if it does not contain personal data. We should be certain that the AI system being used doesn't train on the work that we enter and that it meets the DFE's product safety requirements.

In the example of marking student work, it's important to be transparent about how AI is being used and sure that the AI tool is fit for purpose. It must always be used with human oversight and not used to replace the important human feedback we know can be very effective in improving progress. Educators' organisations and AI developers must be aware of both requirements to ensure compliance with legal and ethical standards when using AI tools.

It's also important to only use AI systems approved by your setting.

In the example we're about to hear from Scott Hayden, his college provided access to an AI system for students. They did this very carefully and ensuring that it was a closed system that doesn't train on the data or prompts given by users. It also has appropriate monitoring, filtering and reporting built in with regard to the DFE's product safety expectations and to Keeping Children Safe in Education guidance. They also provided key training for the students to help them get the best out of it and use it safely.

Scott Hayden, Head of Teaching, Learning and Digital, Basingstoke College of Technology:

“So the way we've considered the potential drawbacks of AI is to be very transparent with our policy when learners start our courses at the college. So they are presented with the policy, dos and don'ts and they agree to it in enrolments and they have it made very clear to them as part of their AI module they do in week one as part of their induction, how to use AI the right way, and also what constitutes using AI the wrong way. We try to make that as unambiguous as possible.

“We also state very clearly that it is against our policy to counterfeit human likeness. We have zero tolerance for counterfeiting human likeness. So what I mean by that is to take somebody and to counterfeit text, audio, image or video of somebody else without their consent. We have zero tolerance for that because we've seen horrific instances of this being done. And with the rise of deep fakes and other uses of AI that are less than nice we need to be very clear about the fact that we use AI to help us learn, use it the right way, and we need to show them very clearly and state very clearly what is unacceptable.”

Presenter: It's important to remember to follow your setting’s guidance on safeguarding and online safety. Educators and leaders are reminded that the Department for Education's Keeping Children Safe in Education’s statutory guidance provide schools and colleges with information on what they need to do to protect pupils and students online, their responsibilities with regards to limiting children's exposure to risks from the school's or college's IT system, how to review and strengthen their cybersecurity – generative AI could be used to increase the sophistication and credibility of attacks. Again, you can find more about this in the Leadership toolkit.

Video 3 – Evaluating generative AI outputs

Presenter: Welcome to module three, video three from the safe and effective use of AI in education online resources.

This video explores the importance of careful choices and a critical mindset when using generative AI.

A critical mindset is a kind of thinking in which you question, analyse, interpret, evaluate and make a judgment about something that you read, hear, say or write. We know this is particularly important when using generative AI.

The persuasive tone that many AI systems take can sometimes make it difficult to be objective.

Merve Lapus, Vice President, Education Outreach and Engagement, Common Sense Media:

“And I think those types of considerations are going to be incredibly important as the tools in AI allow for them to create more rapidly and share more widely. In many cases, sometimes making it indistinguishable from something that they might have created over an expanded length of time or that they can create now in five to ten seconds.”

Bukky Yusuf, Deputy Headteacher and Science Lead, Edith Kay School:

“Teachers can critically evaluate the output from generative AI by really being mindful about, first of all, what is it you want the tool to do?

So is it in terms of summarising information?

Is it in terms of ensuring that resources are presented in an engaging way?

Having the framework in mind about what you want it to do and then checking, has it met that particular criteria?

But also, more essentially, ensuring that the information it provides is accurate as well. And it's appropriate to the learning needs of the students that they teach.”

Dr Neelam Palmer, Director of Digital Learning and Education, AISL Harrow Schools in Asia:

“I think one of the things that we don't talk about enough is AI literacy. There is AI in teaching and learning and keeping our children safe, and extending their understanding and use of it. But then there's this whole other area of AI literacy. We have actually developed curriculum plans in terms of bringing AI lessons into teaching and learning from a more technical point of view. Whereas our teachers know things about being safe,

they know about cyber security, they know about innovation and multimedia within a computing curriculum. They don't know enough about AI literacy.

“Those components of understanding the history of AI, understanding what it means to work with narrow AI, wide AI, and what AI is composed of, what the implications of using AI in various scenarios are – we're not focusing enough on that, and I think that conversation needs to emerge in this space.”

Peter Reeves, Head of Digital Education Training and IT, Cavendish Education Group:

“If we're looking at specific neurodiversity needs, AI struggles sometimes. A lot of the content that's out there for AI is generated by neurotypical mindsets. It's generated with neurotypical data. It's generated with neurotypical examples. If you're looking at a quick resource, I can see its benefit, but if you're really looking to fine-tune that with SEN expertise, that's where the human comes in. That's where AI gives us that starting point and then we develop that resource. I haven't found a resource yet that really meets that need when you really drill down into it.”

Presenter: AI can be an incredibly useful tool for saving time, streamlining tasks and improving efficiency in areas such as policy writing, lesson planning and letter writing, it can quickly generate drafts, summarise key points from lengthy documents and compare existing policies against new legislation. This can free up valuable time for school leaders and staff, allowing them to focus on other priorities.

However, while AI can take on much of the heavy lifting, there are moments where outsourcing too much of the thinking process can be risky. Some tasks require deep engagement, critical analysis and professional judgment – things that AI cannot replicate. If we allow AI to do all of the work, we risk losing out on valuable learning, understanding and scrutiny that comes from thinking through a problem ourselves.

A clear example of this is when drafting or updating school or college policy. AI can certainly help by structuring a document, suggesting wording and even identifying potential gaps based on new legislation. However, it remains the responsibility of school or college leaders to fully understand the legal requirements and ensure that policy is fit for purpose. Relying solely on AI could mean missing crucial nuances in the law, misinterpreting regulatory updates or failing to tailor a policy to the specific needs of the organisation.

A related challenge is the tendency to anthropomorphise AI, treating it as if it has intentions, reasoning or understanding akin to human intelligence. AI does not think or know in the way people do. It processes patterns and probabilities based on its training data. When AI is framed as a sentient-like entity, users may overtrust its responses or expect it to behave ethically on its own rather than recognising the need for human oversight. This misplaced perception can lead to complacency in evaluating AI-generated content and risks attributing authority to a system that ultimately lacks true comprehension.

Recognising AI as a tool and one that requires scrutiny and contextual understanding helps mitigate these risks.

As we've explored in this video, careful choices and a critical mindset are paramount when interacting with generative AI. It's also important in evaluating whether or not we should choose to use AI for particular tasks. The way many AI systems have been designed to sound persuasive can easily sway our objectivity. By taking a considered approach ourselves, we make it more likely that any output we use from generative AI tools is appropriate for its intended use. By emphasising critical thinking across the curriculum, we empower students and pupils.

Video 4 – Data protection and intellectual property

Presenter: Welcome to module three video four from the safe and effective use of AI in education online resources. In this video, we will explore the importance of data protection and intellectual property when using student work with generative AI, as well as other considerations.

When using generative AI tools, we must remember our responsibilities in protecting the rights of the students and pupils we work with. There are two distinct areas that need to be considered: children's personal data, as protected by UK GDPR, and their intellectual property. These require separate processes of consent and permission, as set out in the DFE's AI policy paper.

Children require specific protection, as they may be less aware of the risks and consequences of generative AI on their rights to their personal data. [You can pause the video now if you need a reminder of what falls into these two categories.]

This means we must raise student and pupil awareness of data protection by being transparent, telling them when and how generative AI systems will use their data in clear and simple language that they will understand, being open about the data protection risks and safeguards involved with the use of generative AI systems, and letting them know what to do if they are unhappy.

The Generative AI and Data Protection in Schools guidance help schools understand AI related risks and biases to personal data. It provides tools and resources to address potential data protection risks and helps schools ensure appropriate levels of protection are in place for students' and pupils' data.

Children have the same rights as adults over their personal data under the UK General Data Protection Regulation or UK GDPR and the Data Protection Act 2018. Under this and alongside other rights, children have the right to: Be provided with a transparent and clear privacy notice, which explains who you are and how their data will be processed. Be given a copy of their personal data. Have inaccurate personal data rectified and incomplete data completed. Exercise the right to be forgotten and have personal data erased.

Where a school decides that consent is an appropriate lawful basis for processing data with AI, the UK GDPR also states that "it is relevant in particular where the data subject has given his or her consent as a child, and is not fully aware of the risks involved by the processing and later wants to remove such personal data, especially on the internet. The data subject should be able to exercise that right, notwithstanding the fact that he or she is no longer a child."

What does this have to do with generative AI? Well, we know the generative AI systems are trained on huge data sets. In some cases, as well as that initial training data, the system continues to be trained through the data that users input into it.

As a separate issue to data protection, it's important to recognise that the intellectual property or IP of work produced by pupils or students belongs to them. Therefore, to avoid infringing intellectual property rights we need to follow the DFE's Generative AI in Education guidance that states:

"Schools and colleges must not allow or cause students' original work to be used to train generative AI models unless they have permission or an exception to copyright applies. Permission would need to be from the student as the copyright owner or the student's parent or legal guardian, if the copyright owner is unable to consent because of being a minor. Exceptions to copyright are limited and settings may wish to take legal advice to ensure they are acting within the law."

Let's take the example of generating feedback on student work. A teacher might paste pupil or student work into a large language model or LLM to generate feedback. The work that the pupil or student has done remains their intellectual property. There are several potential issues in this example. The AI could learn from the data inputted and therefore incorporate the student's work into the model. Without permissions in place, this would infringe copyright law. There is also the need for transparency in any use of AI, especially in such a crucial area as feedback to students. We also need to be particularly careful to check the quality of the feedback and adapt it. Even if we know we are using a safe tool, have permissions in place and are being transparent we must always maintain human oversight of this process.

We strongly recommend as an educator, you only use the generative AI tools provided by your institution. This is because the designated data officers alongside the relevant members of leadership or management should have conducted the necessary checks to ensure appropriate privacy protections are in place for students' and pupils' data, which may include a data protection impact assessment. Free AI tools often don't have these safety considerations in place and require careful consideration, bearing in mind the Department for Education's product safety expectations. Enterprise or professional AI tools incorporated into suites of secure tools for schools and colleges are much more likely to meet these requirements.

The legal landscape is still evolving around the way that large language models work, and we need to consider that copyrighted material may have been used in the training of the model, meaning that there is a potential for the output to resemble or contain copyrighted material. This is known as secondary copyright infringement. We can reduce the risk of intellectual property infringements by avoiding public sharing, for example, on the school or college website, of any content largely created by a generative AI tool. [You can pause the video now if you'd like to look at the statement on this from the DFE's Generative AI and Education policy paper.]

In addition to pupils' or students' work and AI generated work, schools and colleges often use copyrighted materials, and there is a risk in using these with AI systems. For example, when we are creating learning resources, we must not share copyrighted material with an AI tool without the permission of the copyright owner, as this may infringe their intellectual property rights and may be irreversible if the AI system learns from the input.

AI tools are becoming more integrated into systems, and that can make it difficult to identify the data that is being used by the integrated AI, which highlights the importance of schools and colleges regularly reviewing the approved tools. There is more advice for leaders on this in the leadership toolkit.

So to summarise, always remember to only use generative AI tools that are approved and provided by your setting for work purposes. These should meet the DFE's product safety expectations and Keeping Children Safe in Education guidance requirements. Be particularly careful when working with student data and/or intellectual property, only using it in an approved way while respecting pupils' and students' intellectual property and data rights. All tools should have a data protection impact assessment completed. And permissions need to be in place regarding the use of intellectual property. This includes, but is not limited to, pupil or student work.

Personal data should not be entered into AI tools unless you have been advised that it is safe to do so by your data protection officer, and a full impact assessment has taken place.

Be aware that, depending on the generative AI system used, the prompt you may not be private and could be used to train the model. This can be especially the case in free versions of tools, or those provided as a personal account outside of your school or college's a proved suite of digital tools, highlighting the importance of only using the approved tools.

We should always be transparent about how we are using generative AI as educators.

Education settings should ensure they have a lawful basis or consent adhering to UK GDPR for processing data, and they need to be particularly careful when using data in AI tools.

Education settings should also ensure that they have the correct permissions in place to avoid infringing intellectual property.

Video 5 – Academic integrity

Presenter: Welcome to module three, video five from the Safe and Effective Use of AI in Education online resources.

When it comes to exams, schools, colleges and exam boards take breaches of integrity seriously, but integrity must be embedded throughout all learning, including homework and other unsupervised work.

AI presents both opportunities and challenges in education. While it can support learning, it must be used responsibly. The focus must remain on ensuring students and pupils develop their own knowledge and skills rather than relying on AI-generated content.

One of the challenges around academic integrity is that AI-generated content can be difficult to distinguish from student work. Some detection tools claim to identify AI-generated content, but research shows they are unreliable. They can produce false positives, disproportionately affecting students for whom English is an additional language, and often fail to detect well-disguised AI use.

Professional judgement and a clear understanding of students' typical work, rather than relying solely on detection tools, will always be the most effective way to identify potential misuse and avoid incorrectly accusing students.

For exams, JCQ guidance makes it clear that students and pupils must submit work that is entirely their own. If they use AI in any way, it must be appropriately acknowledged. Failure to do so constitutes malpractice and may result in severe consequences, including disqualification.

Teachers and assessors must investigate suspected AI misuse and take appropriate action. The full JCQ guidance can be found in JCQ AI Use in Assessments.

Manual checks are one of the most effective ways to identify AI misuse. The JCQ outlines steps for educators in identifying misuse, keeping watch for inconsistencies in student work, and checking for unusual vocabulary, tone shifts, and unverified references. If plagiarism or AI misuse is suspected, educators should try to confirm the source before reporting through the appropriate channels.

The JCQ's quick guide for students on AI and assessment is a valuable tool to share with students and pupils.

As we become better at discussing risks and spotting intentional and accidental misuse of AI, pupils and students are just as likely to become more adept at using generative AI, and so the likelihood of misuse may also increase.

There is the potential for this to happen during non-examination assessments (NEAs) at Key Stage 4 and 5 level. However, other less formal assessments take place throughout schools and colleges all the time.

These may be done under supervision or away from the classroom at home. Some pupils and students may be equipped with their own devices, whereas others rely on access to classroom-based PCs or laptops. Each context will be different, and how students and pupils access generative AI in support of their work in and out of school or college will require individual interpretation and should be reflected in your organisation's relevant policy.

While plagiarism is the uncredited use of someone else's work, AI misuse can involve over-reliance on AI-generated content in ways that bypass genuine learning.

Schools and colleges must provide clarity on when AI is not to be used, ensuring pupils and students understand their learning should be based on their own efforts.

AI use is not just a concern at Key Stage 4 and 5, and students and pupils of all ages need to understand that their learning should come from their own engagement with the material. Embedding discussions about AI early helps students and pupils recognise that reliance on AI-generated content is not a substitute for genuine academic effort. It is important for students and pupils to understand that AI should not be used in place of independent thought and critical engagement.

Homework presents a particular challenge. AI tools can support educators in preparing learning materials, but it is not something that should be relied on by students and pupils to complete their work. Some may attempt to use AI tools to bypass the learning process, undermining their own development. Others may not yet understand that using AI-generated work as their own is a form of misconduct. Educators should set clear expectations that students' work must be their own, and assignments should be designed to encourage independent thinking.

Tasks that require personal reflection, reasoning, or practical application are more effective at promoting learning and are harder for AI to complete convincingly.

It's important to understand when discussing AI detection tools that they have limitations. False positives can unfairly penalise students, and AI detection tools are not always reliable in distinguishing human and AI-generated work. Schools and colleges should provide guidance on how to approach such cases fairly, ensuring students are not disadvantaged. The focus should remain on ensuring students and pupils develop their own understanding and skills without reliance on AI-generated work.

In written work, you can encourage citation of sources, and in other creative and technical subjects, you can focus on activities that make the process and iterations of the work visible and transparent. Most of all, it's important that students and pupils understand whether the use of AI constitutes cheating. Mock or practice assessments are a good opportunity to practise completing authentic work. If malpractice is suspected, discussing the work with a student can help as part of a holistic approach to investigating.

As new technologies and AI tools emerge, determining their acceptable use will be an ongoing challenge. Making clear the expectations around how much or how little students and pupils can use generative AI for their work, particularly in assessments, is important.

Any access to AI for students needs to be carefully planned, considering the DfE's policy position on the use of AI and its product safety requirements. Access must also align with Keeping Children Safe in Education guidance.

Most freely available tools would likely not be suitable for student use, as they will not have the appropriate safeguards in place, and the model may learn from the prompts and information entered into them.

These conversations need to remain open and ongoing as technologies develop. Always ensure that you follow your setting's policies and guidance on this. By fostering a culture of integrity and setting clear expectations, schools and colleges can ensure AI does not compromise authenticity or bypass the important thinking that is part of the learning process. When students and pupils understand both the risks and ethical concerns of AI, they are better equipped to engage with learning in a way that prioritises their own effort and development.

Department for Education

About this publication:

enquiries https://www.gov.uk/contact-dfe
download www.gov.uk/government/publications

Follow us on X: @educationgovuk
Connect with us on Facebook: facebook.com/educationgovuk