TEACHERMATIC
AVALLAIN GROUP

PEDAGOGY-FIRST AI IN EDUCATION

TeacherMatic SAFE
AI Framework

A pedagogy-first model for the responsible use of AI in feedback and assessment
teachermatic.com

Published for institutional review. We welcome feedback from colleges, universities and sector partners as the TeacherMatic SAFE AI Framework continues to develop.

To discuss the TeacherMatic SAFE AI Framework in your institution, contact the TeacherMatic team via teachermatic.com.

VERSION 2 AUGUST 2026

Contents

01	Executive summary	2
02	Rationale	3
03	Introduction	4
04	Safeguarding data and privacy	5
05	Augmenting professional judgement	7
06	Fairness and inclusion	10
07	Ethical and transparent practice	13
08	Implementing the TeacherMatic SAFE AI Framework: institutional governance	16
09	Implementation aids	18
10	Ongoing development	21
11	Frequently asked questions	22
12	Glossary	25
13	Disclaimer	27

01 Executive summary

The TeacherMatic SAFE AI Framework provides a clear, institution-ready approach to the responsible use of AI, artificial intelligence, in feedback and assessment. It has been developed in response to the rapid emergence of AI-supported feedback across education and the growing need for clear ethical and operational guardrails in this area.

The TeacherMatic SAFE AI Framework ensures that AI-supported feedback remains secure, educator-led, fair and transparently governed. It is built for institutions that want to benefit from AI without compromising academic integrity, professional judgement or learner trust. Institutions should ensure that their policies on student AI use and educator-facing AI tools are coherent and communicated together, so that students understand the distinction between declared, educator-governed AI feedback and undeclared AI use in their own submitted work.

SAFE STANDS FOR

S	Safeguarding data and privacy
A	Augmenting professional judgement
F	Fairness and inclusion
E	Ethical and transparent practice

Unlike general-purpose AI tools that can generate unstructured or autonomous outputs, TeacherMatic's feedback workflows operate within clear pedagogical and governance boundaries. AI-generated feedback is positioned as draft and advisory, never as an automated academic decision.

The TeacherMatic SAFE AI Framework supports a hybrid model of marking where professional expertise remains central and technology reduces administrative burden without compromising academic integrity or learner trust.

02 Rationale

AI-supported feedback is emerging rapidly across the education sector, yet there is still limited clarity around how it should be used responsibly within assessment. Feedback and marking sit at the heart of academic integrity, professional judgement and institutional reputation, so it is understandable that schools, colleges and universities are cautious about introducing AI into such a sensitive area.

For many educators, this remains uncharted territory. Teachers are often navigating AI use independently and may be unsure what is acceptable, where the boundaries lie or whether they may inadvertently be breaching institutional expectations or guidelines. Many institutions also do not have the capacity to develop detailed frameworks quickly and are looking for practical guidance and reassurance.

The case for AI-supported feedback is not only about reducing workload and delays in returning work to students. Prompt, detailed feedback has a direct impact on learning. When feedback is returned weeks after submission, learners have often moved on from the thinking that produced the work, and the opportunity to act on feedback is diminished. AI-supported feedback can close this gap by enabling educators to return higher-quality, criteria-aligned feedback in a shorter timeframe, while retaining final professional judgement.

The TeacherMatic SAFE AI Framework was developed to provide clarity in this space. It offers a structured, principled approach to ensure that AI-supported feedback remains secure, educator-led, fair and transparently governed, and that innovation strengthens rather than compromises learning outcomes, academic integrity and learner trust.

As TeacherMatic introduces its AI-supported feedback generators, including the Advanced Feedback Generator, which enables structured feedback on uploaded student work, it is important that clear ethical guardrails accompany this capability. The TeacherMatic SAFE AI Framework provides a proactive response to the questions institutions are already asking about data protection, academic judgement, transparency and intellectual property.

03 Introduction

The TeacherMatic SAFE AI Framework helps schools, colleges and universities adopt AI tools in ways that protect academic integrity, safeguard data and preserve professional judgement while enabling innovation and reducing workload. The framework reflects both pedagogical best practice and emerging expectations around responsible AI use in assessment.

Across this document, the term AI-supported feedback is used as the primary descriptor. This signals that the educator remains the author of the feedback decision, with AI operating in a supportive role.

Note on regulatory positioning

TeacherMatic's feedback and assessment tools are designed to operate consistently with the principles of the EU Artificial Intelligence Act (2024) UK Department for Education guidance on AI in education, and equivalent emerging international frameworks. The platform's core function is to support educators in generating draft feedback content and to avoid autonomous, high-stakes academic decisions such as final grading or admissions.

By focusing on educator-led workflows, TeacherMatic supports the kind of transparency, oversight and safety expectations these frameworks set out, namely:

- **Transparency.** Users are clearly informed when they are interacting with AI-generated content.
- **Contextual guidance.** Educators are given the framework to interpret and refine AI outputs.
- **Usage records.** Internal logging supports institutional accountability and traceability.
- **Human oversight.** The system architecture prevents autonomous grading, so teachers retain final authority.

Regulatory expectations for AI in assessment are evolving and may classify AI-assisted evaluation as higher risk in some jurisdictions. The TeacherMatic SAFE AI Framework responds to this by treating educator review, transparent disclosure and auditable records as non-negotiable, regardless of the tier under which a given system is ultimately classified.

Awarding body guidance on AI use in assessment varies across organisations and continues to evolve. The framework sets out principles and good practice; institutions must always verify the guidance of each relevant awarding body before implementing AI-supported assessment workflows.

PILLAR 01

S

Safeguarding data and privacy

CORE PRINCIPLE

Student work is sensitive academic and personal data. It must be protected with the highest standards of security and care.

In this framework, the term Safeguarding refers to the protection of student work, identifiers and assessment data within AI-supported workflows. It is distinct from the institution's statutory safeguarding duties for the welfare of children and vulnerable adults, which sit outside the scope of the TeacherMatic SAFE AI Framework.

What this means in practice

TEACHERMATIC

- Uses instances of AI models that do not train on user-submitted data
- Operates within secure, controlled environments supported by internationally recognised security frameworks
- Does not claim ownership of student submissions or generated outputs and does not reuse submitted content beyond the feedback generation task
- Processes uploaded content solely for the purpose of generating feedback on the submitted work
- Aligns with UK GDPR, EU GDPR and equivalent international data protection standards
- Hosts processing on UK-based secure infrastructure with appropriate controls
- Supports secure bulk upload and inline annotation workflows
- Provides role-based access controls and appropriate data retention and deletion processes
- Supports transparency through platform-level logging available to authorised organisational administrators
- Maintains documented incident response procedures, including timely notification to authorised institutional administrators in the event of a data security incident
- Maintains a transparent record of subprocessors and underlying AI model providers, with appropriate contractual safeguards in place and notification of material changes
- Applies retention and deletion policies that continue to apply once a learner has completed their programme or left the institution

- Encourages institutions to follow good practice when uploading work, such as avoiding unnecessary personal identifiers and adhering to institutional data protection policies

Institutional good practice

INSTITUTIONS EMBEDDING THE TEACHERMATIC SAFE AI FRAMEWORK ARE ENCOURAGED TO

- Ensure AI-supported feedback is reflected within assessment and data protection policies
- Confirm retention and deletion procedures are clearly defined and continue to apply when learners complete their programme or leave the institution
- Ensure incident response procedures cover AI-supported feedback workflows alongside other data systems
- Provide clarity to staff about appropriate data handling practices
- Encourage staff to minimise or anonymise unnecessary personal data when uploading work
- Maintain oversight of how student work is processed within AI-supported workflows, including any subprocessors or AI model providers in the chain
- Obtain student consent for AI-supported feedback processes; a tick box on enrolment forms or learner agreements offers a practical and consistent mechanism
- Verify that any AI tools used for feedback processing do not claim ownership of uploaded student work, do not use submitted content to train, update or fine-tune any AI model and do not process the material in any other way that would affect its ownership or privacy
- Seek contractual guarantees from any AI provider around non-training, IP ownership and deletion terms, independent of vendor choice

PILLAR 02

A

Augmenting professional judgement

CORE PRINCIPLE

AI should support educators, not replace them.

What this means in practice

TEACHERMATIC

- Ensures teachers remain in full control of the final output
- Positions AI-generated feedback as draft, advisory or supportive
- Requires educator review, editing and approval before feedback is finalised
- Does not fully automate marking, assign grades or replace the professional judgement of the educator. TeacherMatic generates written, developmental comments on student work; the assignment of marks, grades and pass/fail decisions remains entirely with the educator
- Provides refine and additional-information features that give educators direct control over tone, emphasis and content
- Gives organisation administrators the option to switch on **AI Assisted Marking Acknowledgement**, which requires staff to acknowledge and accept the institution's policies on AI-supported marking before they can use AI-supported feedback
- Gives organisation administrators the option to switch on **Enforce Human in the Loop**, which requires teachers to mark each submission as reviewed before they can bulk download results
- Supports colour-coded annotation so that AI-generated, human-added and hybrid comments can be visually distinguished
- Supports transparency through platform-level logging available to authorised organisational administrators

New in this version: *AI Assisted Marking Acknowledgement and Enforce Human in the Loop are organisation-level settings, switched off by default, that administrators can enable from Organisation Settings.*

The 80 / 20 principle in practice

The TeacherMatic SAFE AI Framework reflects a hybrid marking model in which AI generates approximately 80 per cent of the first-draft feedback and the educator reviews, refines and evaluates the final 20 per cent. The 20 per cent is where professional judgement is applied: interpreting context, moderating tone, challenging AI-generated phrasing and making any final grading decision. The 80 / 20 split is illustrative rather than prescriptive. The underlying principle is that AI handles the heavy lifting and the educator retains authorship of the decision.

In practice, a meaningful educator review can take from a couple of minutes for short-form feedback to ten minutes or more for substantial assignments, depending on assessment type and stakes. Institutions should set local minimum review expectations rather than relying on time alone, and combine these with sampling and moderation to confirm review quality.

When a formal rubric is not available

Formal rubrics are not always available. AI-supported feedback can still be generated where the educator provides an assignment brief, learning outcomes, marking criteria, a mark scheme or other context that anchors the feedback. The less formal the criteria, the more important the educator's professional judgement becomes at the review stage.

When AI and educator judgement diverge

Where the educator's professional judgement differs from the AI-generated draft, the educator's decision is final and must be reflected in the feedback returned to the student. Educator judgement always takes precedence.

Alignment with assessment expectations

Across awarding bodies and regulatory guidance, a single principle is consistent: AI must not replace human academic judgement in assessment. AI may support feedback generation, but final academic judgement and marking decisions remain with qualified educators.

Institutional good practice

INSTITUTIONS EMBEDDING THE TEACHERMATIC SAFE AI FRAMEWORK ARE ENCOURAGED TO

- Reinforce that final academic decisions remain with qualified educators
- Provide guidance on responsible AI-supported marking practices
- Integrate AI use into moderation and quality assurance processes
- Avoid policies that promote automated grading without human oversight
- Put in place verifiable human review mechanisms, for example a minimum review standard, or a workflow-level record that an educator has opened and reviewed each piece of feedback before release
- Decide whether to enable Enforce Human in the Loop and AI Assisted Marking Acknowledgement at organisation level, so review and policy sign-off are enforced at the platform level
- Use colour coding or equivalent annotation conventions to distinguish AI-generated comments from human-added and hybrid comments within annotated student work
- Invest in staff induction and ongoing support so educators engage meaningfully with AI-generated feedback rather than releasing it without substantive review

IN PRACTICE LARGE UNDERGRADUATE COHORT, FIRST-YEAR MODULE

An engineering lecturer used AI-supported feedback on a group design assignment with four tasks. Every feedback sheet was reviewed against their own professional judgement before release. A short statement on each feedback sheet made clear that the lecturer had marked the work and used AI to support feedback generation. No students raised concerns.

PILLAR 03

F

Fairness and inclusion

CORE PRINCIPLE

AI-supported feedback should promote consistency, inclusivity and criteria alignment.

What this means in practice

TEACHERMATIC

- Aligns feedback with rubrics, learning outcomes and assessment criteria
- Applies multi-layered content filtering to reduce harmful, biased or inappropriate outputs
- Applies prompt-level and model-level measures to mitigate known biases carried by underlying large language models
- Encourages educator review to identify and address any unintended bias in individual outputs
- Supports consistency across markers, cohorts and departments by applying the same criteria regardless of cohort size or marking sequence
- Reduces markers of fatigue-related and relational bias that can emerge in extended marking sessions
- Allows educators to refine the phrasing, tone, emphasis and accessibility of generated feedback so that it reflects institutional, disciplinary and learner-specific needs

SEND (Special Educational Needs and Disabilities) considerations, accessibility of feedback formats and inclusive design sit with the institution and educator within the TeacherMatic SAFE AI Framework. The platform supports inclusive practice by enabling educators to adapt generated feedback for tone, language and accessibility before release.

A note on model-level bias

A reasonable concern raised by academics and senior leaders is that large language models carry their own biases from the data they were trained on. The TeacherMatic SAFE AI Framework does not dismiss this concern. This can shape the tone, examples and reasoning patterns that surface in AI-generated feedback, in ways that may not always serve every learner equally.

TeacherMatic mitigates this through several mechanisms, including prompt engineering that grounds outputs in the uploaded brief, rubric, learning outcomes or other assessment criteria provided by the educator; content filtering applied on top of model outputs; and the requirement for educator review before any feedback is released.

Equitable access and cohort fairness

AI-supported feedback should support an equitable learner experience across cohorts. Where it is used in some programmes or cohorts and not others, institutions should monitor whether learner experience and outcomes remain comparable, and avoid creating an implicit two-tier feedback culture.

Institutional good practice

INSTITUTIONS EMBEDDING THE TEACHERMATIC SAFE AI FRAMEWORK ARE ENCOURAGED TO

- Combine AI-supported feedback with the rubrics, marking guidance and essential assessment information provided by each relevant awarding body
- Monitor consistency of AI-supported feedback across programmes and departments as part of quality assurance
- Keep inclusive assessment policies central to how AI is used, including SEND, language, accessibility and cultural context considerations
- Ensure that feedback formats are compatible with assistive technologies and inclusive of learners using screen readers, alternative input devices and translation tools
- Provide staff development on bias awareness, inclusive feedback practice and critical review of AI outputs
- Sample a proportion of AI-supported feedback through moderation to confirm consistency with human professional judgement

PILLAR 04

E

Ethical and transparent practice

CORE PRINCIPLE

Students and staff should understand how AI is used in feedback and assessment.

What this means in practice

TEACHERMATIC

- Encourages clear communication with students about AI-supported feedback
- Supports institutional transparency statements and policy alignment
- Avoids hidden or autonomous marking practices
- Embeds documented risk assessments and governance processes
- Provides clear mechanisms for monitoring, reporting and escalation
- Provides platform-level logging available to authorised organisational administrators, so institutions can evidence where AI has been used
- Supports disclosure to external examiners and awarding bodies that AI-supported feedback tools are used within the institution's assessment workflow

Framing AI-supported feedback to students

Evidence from sector pilots suggests that student acceptance is shaped more by how AI-supported feedback is framed than by whether AI is used at all. Framing the use of AI around consistency of marking, timeliness of return, the depth and volume of feedback and the reduction of marker fatigue bias tends to be received more positively.

Transparency is particularly important where students are expected to declare their own AI use. Institutions should be prepared to explain why AI-supported feedback by educators is treated differently from unattributed AI use by students. The honest answer is that educator use is declared, bounded, reviewed and recorded, and the TeacherMatic SAFE AI Framework is what makes that visible.

Transparency with staff

Transparency applies to staff as well as students. Hidden or undisclosed AI use, even at institutional level, undermines the trust the framework relies on.

Learner appeals and complaints

AI-supported feedback should not introduce new categories of academic appeal. Existing institutional appeal and complaint procedures apply equally to feedback generated with AI support, with no special carve-out.

Institutional good practice

INSTITUTIONS EMBEDDING THE TEACHERMATIC SAFE AI FRAMEWORK ARE ENCOURAGED TO

- Provide student-facing statements explaining how AI-supported feedback is used
- Ensure AI use aligns with academic integrity regulations
- Communicate clearly with staff about expectations and boundaries, including where AI is used in moderation and quality assurance workflows
- Confirm that institutional academic appeal and complaint procedures apply equally to feedback generated with AI support, with no special carve-out
- Maintain a documented record of where and how AI has been used within assessment workflows
- Exercise significant caution with AI-generated plagiarism or academic integrity detection tools, which currently produce high rates of false positives

08 Implementing the TeacherMatic SAFE AI Framework: institutional governance

Compliance with the TeacherMatic SAFE AI Framework is an institutional commitment, not an individual one. It requires organisational readiness, clear governance and active senior leadership support.

Institutions planning to embed AI-supported feedback are strongly encouraged to establish a cross-functional working group bringing together senior leaders, middle managers, teaching staff, business support staff and student representatives. This group should address four foundational areas before rollout begins: permissions, policies, procedures and people.

Permissions

Students should be informed that AI tools may be used to support feedback generation and should provide consent. For institutions with upcoming intake cycles, the most practical approach is a tick box on enrolment forms or learner agreements.

Policies

Existing institutional policies may not yet fully address the use of AI in assessment and feedback. Any institutional approach should reflect the hybrid marking model, the human-in-the-loop requirement and the institution's position on which assignment types and awarding bodies are in scope.

Procedures

Staff need practical guidance on how to use AI-supported feedback tools responsibly, including the minimum review standard required before feedback is released, whether Enforce Human in the Loop and AI Assisted Marking Acknowledgement should be switched on at organisation level, and how AI use should be recorded for quality assurance purposes.

People

Technology adoption succeeds or fails on culture. Practical measures include induction sessions for new users, short peer-observation slots, communities of practice, and regular feedback loops into the institutional working group.

Staged adoption: formative first

Sector pilot evidence suggests a clear starting point. Formative feedback, including draft review, practice essays and reflective writing, is the lowest-risk entry point for AI-supported feedback. Institutions are encouraged to begin with formative use cases, build staff and student confidence,

and extend into summative contexts only when governance, moderation and staff capability are demonstrably in place.

09 Implementation aids

The resources below are provided as indicative starting points. Institutions should adapt them to their own context, policies and awarding body requirements.

The educator-led workflow

01 UPLOAD	02 GENERATE	03 REVIEW	04 RELEASE	05 RECORD
Brief, rubric or assessment criteria uploaded by educator	AI generates draft feedback aligned to criteria	Educator reviews, refines and approves	Feedback released with AI-use declaration	Use logged for quality assurance and audit

The TeacherMatic SAFE AI Framework at a glance

Principle	What to demonstrate
S Safeguarding	Secure processing, no model training on uploaded work, clear retention and deletion, IP retained by institution or learner
A Augmenting	Educator-led workflow; configurable Enforce Human in the Loop and AI Assisted Marking Acknowledgement controls; AI positioned as draft and advisory
F Fairness	Rubric-aligned outputs, inclusive practice, mitigation of fatigue, relational and model-level bias
E Ethics	Student-facing statements, staff guidance, documented record of AI use, honest framing

Institutional self-assessment checklist

Use the prompts below to gauge readiness. Each item can be rated as in place, in progress or not yet started.

Permissions

- Student consent mechanism is agreed and in place
- Student-facing statement on AI-supported feedback is drafted

Policies

- AI use is reflected in assessment, academic integrity and data protection policies
- Position on formative and summative use is clear
- Awarding bodies in scope have been checked individually

Procedures

- Minimum educator review standard is defined and communicated
- Decision made on whether to enable Enforce Human in the Loop at organisation level
- Decision made on whether to enable AI Assisted Marking Acknowledgement at organisation level
- Record of AI-supported feedback use is maintained for quality assurance

People

- Induction and CPD are available to all staff with access to the tool
- Programme leaders are equipped to have direct conversations with cohorts

10 Ongoing development

The TeacherMatic SAFE AI Framework will continue to evolve as sector practice develops, informed by institutional feedback, regulatory guidance and ongoing research into responsible AI use in education.

TeacherMatic's ongoing participation in sector pilots, including the JISC AI assessment pilot, is central to this development process. As comparative feedback, BTEC assessment, portfolio-based grading and VLE integration capabilities are developed, the governance and ethical considerations that accompany them will be reflected in future versions of this framework.

Institutions using TeacherMatic SAFE AI Framework-aligned tools are encouraged to share their own experiences and governance approaches with TeacherMatic, contributing to a growing body of shared practice across the sector.

11 Frequently asked questions

Concise responses to the questions institutions most often raise about AI-supported feedback and the TeacherMatic SAFE AI Framework.

What is AI-supported feedback?

AI-supported feedback is feedback drafted with the help of an AI tool and then reviewed, refined and approved by a qualified educator before it is released to a learner.

Who decides the final mark?

The educator. TeacherMatic SAFE AI Framework-aligned tools do not determine grades. Final academic judgement remains with qualified educators.

Where is student data stored, and is it used to train AI models?

Student work is processed within a secure environment, is not used to train third-party AI models, and remains the intellectual property of the learner or the institution.

What if the AI gets it wrong?

Educator review exists for exactly this reason. Institutions may also choose to switch on Enforce Human in the Loop at organisation level so that review before release is enforced by the platform.

Do students need to consent?

Students should be informed that AI tools may be used to support feedback generation and should provide consent, typically via a tick box on enrolment forms or learner agreements.

How do we evidence governance and oversight?

Through documented policies, recorded consent, platform-level logging, sampled moderation, and a documented record of where AI has been used. Where enabled, the Enforce Human in the Loop and AI Assisted Marking Acknowledgement settings add a platform-level record of educator review and policy acknowledgement.

Who in the institution owns this?

Most successful adopters establish a cross-functional working group bringing together senior leadership, teaching staff, quality, data protection and student representatives.

12 Glossary

Plain-language definitions of the terms used in this framework.

Advanced Feedback Generator TeacherMatic's capability for generating structured feedback on uploaded student work.

AI Assisted Marking Acknowledgement An organisation-level TeacherMatic setting, switched off by default, that requires staff to acknowledge and accept the institution's policies on AI-supported marking before use.

AI-supported feedback Feedback drafted with AI assistance and reviewed, refined and approved by an educator before release.

Enforce Human in the Loop An organisation-level TeacherMatic setting, switched off by default, that requires teachers to mark each submission as reviewed before bulk downloading results.

GDPR General Data Protection Regulation. Sets out rules for the processing of personal data in the UK and EU.

Grading The assignment of a final mark or grade to capture performance at a defined point in a programme.

Rubric A structured scheme used to mark student work against agreed criteria and standards.

SEND Special Educational Needs and Disabilities, a UK term covering learners with additional support needs.

TeacherMatic SAFE AI Framework Safeguarding Data and Privacy, Augmenting Professional Judgement, Fairness and Inclusion, Ethical and Transparent Practice. Always paired with TeacherMatic; SAFE is not used on its own.

VLE Virtual Learning Environment, the digital platform an institution uses to deliver and manage learning.

13 Disclaimer

The TeacherMatic SAFE AI Framework is intended as practical guidance to support institutions in adopting AI-supported feedback responsibly. It does not constitute legal advice. Institutions with specific regulatory or compliance questions are encouraged to seek independent professional advice appropriate to their context.

Each institution remains responsible for its own policies, assessment practices and compliance obligations. TeacherMatic makes no representation that following this framework ensures compliance with any specific legal requirement, awarding body rule or regulatory standard.

This framework reflects the state of sector knowledge and regulatory guidance at the time of publication. It will be reviewed and updated as AI governance in education evolves.

THE TEACHERMATIC SAFE AI FRAMEWORK

Pedagogy-first AI for education.

teachermatic.com