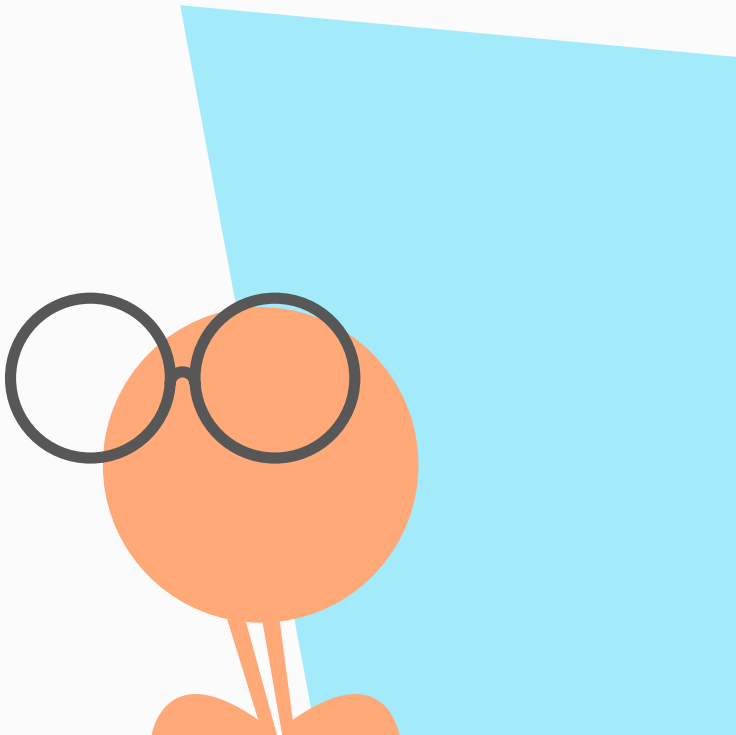
graide

From Policy to Practice

What DfE Guidance & the EU AI Act Mean for Assessment



Introduction

Artificial intelligence is transforming the education landscape. From streamlining marking to generating feedback and tailoring learning journeys, AI promises greater efficiency and personalisation. However, a growing regulatory focus is now tempering this promise.

The UK Department for Education (DfE) and the EU have both released regulations to abide by. The DfE's position on AI in education and the newly enacted EU AI Act will collectively signal a shift from experimentation to accountability.

In this paper, we explore how these evolving regulations apply specifically to assessment, a high-risk domain in educational AI use. Based on Graide's expert-led webinar featuring **Dr Gray Mytton** (NCFE), **Pat Coates** (eAssessment Association), and **Dr. Manjinder Kainth** (CEO @Graide), we break down what policy means in practice alongside the operational, legal and pedagogical implications for assessment providers, institutions and technology developers.

Key Takeaways

Assessment is a high-risk AI application under both DfE guidance and the EU AI Act. This includes not just summative exams but also formative and diagnostic activities.

Safe and effective AI use requires new governance models, not just tools. Institutions should reflect on their internal policies, training and oversight structures.

Intellectual property is a major legal gap. Student work used for AI training cannot rely solely on GDPR compliance.

Transparency and explainability are essential. Black-box AI models may offer efficiency but lack the reliability needed for fair, replicable marking.

Graide's approach of supervised machine learning with human-in-the-loop design is directly aligned with the compliance requirements emerging from both UK and EU regulation.

What Is AI in Assessment?

AI used in assessment typically aims to automate or support activities such as

- Grading student work
- Generating feedback
- Flagging learning gaps or plagiarism
- Informing learning pathways

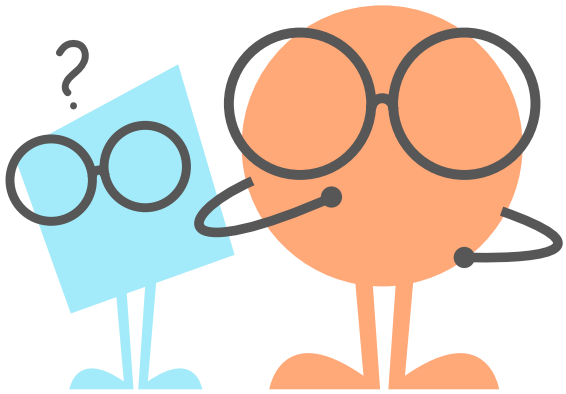
AI operates through three core layers:

1. **Training Data:** Pairs of inputs (e.g., student work) and outputs (e.g., grades, feedback)
2. **Learning Method:** Rules, algorithms, or neural networks applied to identify relationships
3. **Model Use:** Applying the trained system to new inputs to generate predictions or outputs

There are two common types of AI models:

- **Classification AI** (e.g. AI assisted grading)
- **Generative AI** (e.g. producing feedback text)

Each carries different levels of risk and interpretability, with generative AI typically presenting greater challenges in explainability and governance.



DfE Guidance: A Framework for Responsible Use

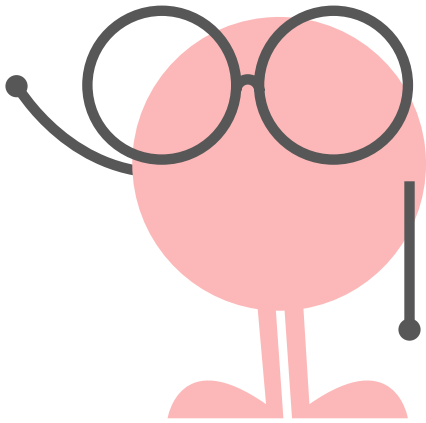
The UK Department for Education has identified some core priorities when it comes to using AI in education:

1. **Safe and Effective Use -** Educators must remain responsible for AI outputs. While productivity tools carry lower risk, anything involving judgement, especially in assessment requires heightened scrutiny.
2. **Responsible Deployment -** Institutions must introduce clear guidelines for when, how and why AI is used in assessment and ensure educators retain agency.
3. **Intellectual Property -** Student submissions used to train AI models remain their intellectual property. Using them without explicit consent creates legal risk even if GDPR provisions have been followed.
4. **Use in Formal Assessments -** Any AI system used in high-stakes decisions (e.g., exam grading) must meet the highest bar of oversight and accountability.

> “Assessment should benefit from AI, but policy must acknowledge the imperfections of current systems and allow space to innovate.”
> **– Coates, CEO @ e-Assessment Association**

The EU AI Act: Global Implications for Education

Though passed by the European Union, the AI Act’s risk-based framework is already influencing UK and global thinking. Its relevance lies in how it categorises education-related AI as high-risk due to its potential impact on learners’ futures.



Article	Focus
6	Classifying high-risk AI (includes summative assessment, admissions decisions and course placements)
8	Compliance obligations for high-risk systems
10	Data governance and consent requirements
13	Transparency obligations to users
14	Human oversight and ability to override AI decisions

The implications are broad. Any AI system contributing to learner outcomes, even if not the final decision maker, must now be governed with the same rigour as traditional assessments.

> “Decisions made by AI about learners, whether course placement or certification, are high risk by nature.”
> **– Dr Gray Mytton, Assessment Innovation Manager @ NCFE**

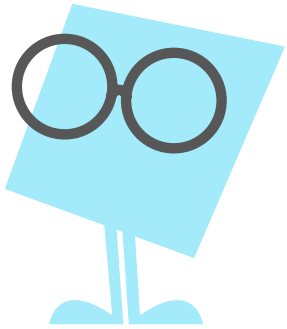
Data, Consent and IP: The Legal Grey Areas

One of the most pressing challenges identified is the ambiguity around data usage for training AI systems. While GDPR addresses data privacy, it does not fully cover:

- Ownership of learner work
- Reuse of assessments or submissions in training datasets
- Responsibility for obtaining IP permissions

Institutions must begin implementing:

- Clear opt-in/opt-out models
- Plain English consent explanations
- Audit trails of training data sources



> The difficulty isn't in writing the explanation; it's building a system that actually respects opt-outs."
> **– Dr Gray Mytton, Assessment Innovation Manager @ NCFE**

Academic Integrity and AI Misuse

The DfE provides guidance on misuse of generative AI by students, including:

- Submitting AI-generated work without attribution
- Fabricating sources
- Overrelying on AI rather than demonstrating independent understanding

Their recommendation:

- Allow AI use but require transparent referencing
- Encourage reflective statements explaining how AI was used
- Shift assessment design toward analysis, reasoning and synthesis

This approach acknowledges AI's ubiquity while preserving pedagogical integrity.

Choosing the Right AI Assessment Tool

Graide's framework (as presented in the webinar) compares five popular AI approaches:

Method	Explainable?	Bias Control	Consistency	Data Volume
Rules-Based	High	High	High	Medium
Supervised ML (Graide)	High	High	High	Low
Neural Networks	Low	Mixed	Medium	High
Fine-Tuned LLM	Low	Mixed	Low	Low
Generative AI + Markscheme	Low	Low	Low	None

Graide's **supervised machine learning** approach is aligned with both UK and EU standards:

- **Trained only on human-verified data**
- **No use of third party pre-trained internet models**
- **Built-in confidence indicators and human override**
- **Low data volume needed**

Conclusion: From Compliance to Confidence

AI is not going away. But using it responsibly in education, particularly in assessment, requires more than good intentions. It requires:

- **Updated institutional policies**
- **Technical due diligence**
- **Legal clarity on IP**
- **Transparent user practices**
- **Commitment to explainability and fairness**

The institutions that act now will be the ones best prepared to harness AI's benefits ethically, equitably, and effectively. Want to explore compliant, transparent AI-powered assessment?

The institutions that act now will be the ones best prepared to harness AI's benefits ethically, equitably and effectively. Want to explore compliant, transparent AI-powered assessment?

Book a demo with Graide →

graide

www.graide.co.uk
contact@graide.co.uk
Linkedin/graide