

Avallain Position Paper on AI Ethics and Security in Education

September 2026





Introduction

This position paper sets out Avallain's view on the ethical, security and practical use of artificial intelligence and generative AI in education for an international audience. It is written for educators, institutions, publishers and decision-makers who are not specialists in law or technology but are responsible for adopting AI-enabled tools in a thoughtful and responsible way.

Avallain designs digital learning platforms that support teaching, learning and educational publishing. Across Avallain Author, Avallain Magnet and TeacherMatic, AI is used to assist professional users, not to replace human judgement, pedagogical expertise or institutional responsibility. Avallain's approach is guided by European principles for trustworthy, human-centred AI, including those set out in the EU AI Act, the EU Ethics Guidelines for Trustworthy AI and data protection frameworks such as GDPR, which Avallain uses as a design and governance baseline.



Scope and Purpose of AI Use

Avallain's AI tools and AI-powered features are designed to support clearly defined educational needs. These capabilities assist publishers, teachers, authors, editors and educational organisations with tasks such as drafting content, generating ideas, summarising material, creating practice activities and improving productivity in content creation and planning. In Avallain Author, AI-powered features support professional content creation workflows for publishers and institutions. In TeacherMatic, AI helps educators prepare lessons, resources, feedback and assessments more efficiently. In Avallain Magnet, AI-integrated features support teachers across a range of learning instances where clients explicitly enable those tools.

Avallain's AI tools are not designed to replace teachers, to make autonomous educational decisions or to evaluate learners without human involvement. They are not intended to determine final grades, progression, certification or learner outcomes. They are not intended to function as general-purpose tutoring systems or substitutes for professional judgement.

Avallain's Role in the AI Value Chain

Avallain acts as an AI system provider. This means that Avallain designs, integrates and operates educational applications built on third-party AI models, while not developing or training the underlying foundation or general-purpose models itself. Avallain is responsible for how AI capabilities are embedded into its products, how they are presented to users and what safeguards and controls are applied at the application level.

Responsibility for AI use is shared across the AI value chain. AI model providers are responsible for the behaviour and limitations of their models. Avallain is responsible for product design, risk mitigation, user guidance and application-level safeguards. Clients and institutions are responsible for deciding whether and how to enable AI features, and for ensuring appropriate use within their organisational and regulatory context. Educators and professional users remain responsible for reviewing, adapting and validating AI-generated outputs before they are used in teaching or learning.

Risk Classification under the EU AI Act

Educational AI systems can fall into high-risk categories when they make or influence decisions that significantly impact learners' rights or opportunities, particularly when placed on the European market or when generating outputs used within it. Avallain's AI-powered features are deliberately designed to avoid such high-risk use cases, establishing an ethical baseline for clients both within and outside the European Union.

Avallain's design ensures that AI outputs are mediated by qualified professionals, or, at minimum, supervised when mediation is impossible. The absence of automated educational decision-making means that AI does not directly determine final grades, assessment outcomes or learner progression without an opportunity for the educator to intervene. These design choices significantly reduce both regulatory and ethical risks by ensuring that AI remains a support tool rather than a decision-maker.

Human Oversight and Professional Responsibility

Human oversight is essential in educational AI because education is a context where judgement, empathy, contextual understanding and accountability matter. AI systems do not understand learners, curricula or classroom realities in the way educators do.

Avallain discourages fully automated grading or evaluation as such practices risk oversimplification, bias and lack of transparency. Automated outputs can appear authoritative while being incomplete, incorrect or misaligned with pedagogical intent. By keeping educators firmly in the loop, Avallain supports responsible use where AI augments professional work rather than replacing it.

Transparency, Explicability and Comprehensibility

Avallain aims to ensure that users are clearly informed when AI is being used. AI features are presented as assistive tools, not as authoritative sources of truth. Users are encouraged to understand that AI-generated output is probabilistic, based on patterns in data and may contain errors, omissions or unintended bias.

For non-technical users, the most important understanding is practical rather than technical. AI can be helpful, but it does not think, reason or verify facts in a human sense. Outputs must be reviewed, edited and contextualised by professionals before use wherever that is possible.

Ethics, Bias and Content Safety

Generative AI introduces ethical risks in education, including the possibility of biased or discriminatory content, culturally insensitive examples, hallucinated facts and reinforcement of dominant perspectives at the expense of diversity and inclusion.

Avallain applies layered safeguards to mitigate these risks. These include infrastructure and model-level content filtering provided by AI model partners. At the application level, Avallain has designed additional controls, such as the Ethics Filter, which strengthen prevention of ethically problematic responses and support safer use of generative AI in educational settings. Avallain explicitly acknowledges that safeguards reduce risk but do not eliminate it. Residual risk is an inherent characteristic of generative AI. Transparency about this limitation is a core part of trustworthy AI design.

Information Security

Alongside ethics and privacy, Avallain considers information security a foundational pillar of trustworthy AI use in education. At a principles level, this means designing AI-enabled systems so that educational content, institutional data and user interactions are handled in a manner that supports reliability, resilience and appropriate access. As an AI system provider, Avallain focuses on responsible system design, clear boundaries of use and shared responsibility across the AI value chain, recognising that information security underpins trust, supports human oversight and aligns with European expectations for human-centred and dependable digital systems.

Data Protection and Privacy

Avallain acts as data controller by default for its products, meaning it determines the purposes and means of processing personal data. In some enterprise contexts, Avallain may act as a data processor where this is explicitly agreed with a client under a data processing agreement.

Data minimisation principles guide how AI features are designed and used. Personal data is processed only where necessary for the intended educational purpose. By default, customer content is not used to train general-purpose AI models. Customer content is used only within the scope of providing the contracted services and is not exposed to other customers or used in any way to inform their generations.

For a detailed description of the technical measures that enforce these commitments, including data sovereignty controls, PII masking, encryption in transit and zero-retention guarantees, see the ['Avallain AI Trust Model: A Personal Data Protection Framework for Generative AI Workflows'](#).

Protection of Children and Learners

TeacherMatic is designed for educators rather than for direct use by students. Educators retain control over how AI-generated resources are reviewed and used with learners. This design choice significantly reduces risks related to child safety, exposure to inappropriate content and unmediated interaction between learners and generative AI.

Across the broader Avallain product suite, institutions retain responsibility for configuring access, supervising use and ensuring that AI-enabled tools are deployed in line with child protection, safeguarding and educational policies. AI does not remove or replace institutional duties of care.

Reliability, Limitations and Residual Risk

AI-generated educational content cannot be assumed to be fully accurate or complete. Generative models may produce plausible-sounding but incorrect information, outdated references or content that does not align with specific curricular requirements.

For this reason, Avallain positions AI output as a starting point rather than a finished product. Professional review, adaptation and validation remain essential steps before content is used with learners.

Continuous Monitoring and Improvement

Avallain continuously evaluates AI-related risks and safeguards through ongoing research, systematic testing and close collaboration with internal product teams, as well as external pilots with educators and schools. This work is further strengthened by partnerships with leading institutions and experts in digital education and AI, including the Centre for Research in Digital Education at the University of Edinburgh, as well as recognised scholars such as Professor Rose Luckin and Professor John Traxler.

Ethical considerations, user feedback and evolving regulatory expectations are systematically incorporated into ongoing improvements to product design, user guidance and published recommendations for the wider educational community. This is an iterative process. As AI technologies and educational practices evolve, so too must the safeguards and assumptions that govern their use.

Guidance and Conclusion

Avallain encourages institutions and educators to adopt AI tools thoughtfully. This includes enabling AI features only where they support clear pedagogical goals, maintaining human oversight at all times, training staff to understand AI limitations and embedding AI use within existing ethical and professional frameworks.

Avallain's position aligns with the EU's vision of trustworthy AI for education and uses these principles as a global reference point across Avallain's international operations, reflecting established approaches for ethical, human-centred digital systems and robust safeguards. By focusing on support rather than substitution, transparency rather than opacity and shared responsibility rather than automation, Avallain aims to contribute to the responsible use of AI in ways that respect learners and empower educators.