Cuadernos de RES PUBLICA en derecho y criminología
Nº en prensa ISSN: 2990-0697 DOI: 10.46661/respublica.13580.

Prevention and criminal offences in university education in the face of artificial intelligence. Regulatory framework, institutional response and challenges of the European AI Act

Prevención e ilícitos penales en la educación universitaria ante la inteligencia artificial. Marco normativo, respuesta institucional y desafíos del AI Act europeo

César Augusto Giner Alegría
UCAM, Universidad Católica de Murcia. Murcia (España)
caginer@ucam.edu
https://orcid.org/0000-0002-9743-7414

Abstract

The integration of artificial intelligence (AI) into university education has profoundly transformed teaching and assessment methods, while simultaneously opening new avenues for the commission of criminal offences. This article analyses the types of offences facilitated by AI in academia — automated plagiarism, identity theft, document forgery, fraud in automated assessments and production of illegal content — and examines in depth the applicable regulatory framework, with particular attention to Regulation (EU) 2024/1689 on Artificial Intelligence (AI Act) and its direct implications for higher education institutions. It also analyses the most effective prevention and institutional response strategies, identifying the existing gaps between technological reality and current normative and organisational capacities. The methodology is qualitative and exploratory, based on systematic literature review, documentary analysis of the AI Act and semi-structured expert interviews. The findings reveal a progressive increase in these offences, the insufficiency of the pre-existing legal framework and the urgent need for robust institutional policies that articulate prevention, digital ethics training and academic compliance mechanisms. This article aims to systematise the landscape of AI-facilitated criminal offences in the Spanish and European university context, analyse in depth the implications of the AI Act for educational institutions, and propose a model of institutional response that integrates prevention, detection, sanction and a culture of integrity. To this end, an interdisciplinary perspective is adopted that integrates criminal law, criminology, applied ethics and educational technology studies.

Keywords: artificial intelligence, university education, AI Act, institutional respons, academic integrity, compliance.

How to cite this work: Giner Alegría. César Augusto. *(2026). Prevention and criminal offences in university education in the face of artificial intelligence: Regulatory framework, institutional response and challenges of the European AI Act. Cuadernos de RES PUBLICA en derecho y criminología, (8), 01–15.* https://doi.org/10.46661/respublica.13580.

Recepción: 06.06.2026 **Aceptación:** 11.06.2026 **Publicación:** en prensa

Cuadernos de RES PUBLICA en derecho y criminología
Nº en prensa ISSN: 2990-0697 DOI:*10.46661*/respublica.13580.

Prevención e ilícitos penales en la educación universitaria ante la inteligencia artificial: Marco normativo, respuesta institucional y desafíos del AI Act europeo

Prevention and criminal offences in university education in the face of artificial intelligence: Regulatory framework, institutional response and challenges of the European AI Act

César Augusto Giner Alegría
UCAM, Universidad Católica de Murcia. Murcia (España)
caginer@ucam.edu
https://orcid.org/0000-0002-9743-7414

Resumen

La integración de la inteligencia artificial (IA) en la educación universitaria ha transformado profundamente los métodos de enseñanza y evaluación, al tiempo que ha abierto nuevas vías para la comisión de delitos. Este artículo analiza los tipos de delitos facilitados por la IA en el ámbito académico —plagio automatizado, suplantación de identidad, falsificación de documentos, fraude en evaluaciones automatizadas y producción de contenidos ilegales— y examina en profundidad el marco normativo aplicable, prestando especial atención al Reglamento (UE) 2024/1689 sobre inteligencia artificial (Ley de IA) y sus implicaciones directas para las instituciones de educación superior. También analiza las estrategias de prevención y respuesta institucional más eficaces, identificando las brechas existentes entre la realidad tecnológica y las capacidades normativas y organizativas actuales. La metodología es cualitativa y exploratoria, basada en una revisión sistemática de la literatura, el análisis documental de la Ley de IA y entrevistas semiestructuradas a expertos. Los resultados revelan un aumento progresivo de estos delitos, la insuficiencia del marco jurídico preexistente y la urgente necesidad de políticas institucionales sólidas que articulen la prevención, la formación en ética digital y los mecanismos de cumplimiento académico. El presente artículo tiene por objeto sistematizar el panorama de los delitos facilitados por la inteligencia artificial en el contexto universitario español y europeo, analizar en profundidad las implicaciones de la Ley de Inteligencia Artificial para las instituciones educativas y proponer un modelo de respuesta institucional que integre la prevención, la detección, la sanción y una cultura de la integridad. Para ello, se adopta una perspectiva interdisciplinar que integra el derecho penal, la criminología, la ética aplicada y los estudios sobre tecnología educativa.

Palabras clave: inteligencia artificial, educación universitaria, AI Act, respuesta institucional, integridad académica, compliance.

Cómo citar este trabajo: Giner Alegría. César Augusto. *(2026).* Prevención e ilícitos penales en la educación universitaria ante la inteligencia artificial. Marco normativo, respuesta institucional y desafíos del AI Act europeo*. Cuadernos de RES PUBLICA en derecho y criminología, (8), 01–15.* https://doi.org/10.46661/respublica.13580.

Recepción: 06.06.2026 **Aceptación:** 11.06.2026 **Publicación:** en prensa

1. Introduction

Artificial intelligence (AI) has burst onto the contemporary scene as a transformative force across all sectors of society, and higher education has been no exception. Over the past decade, universities have progressively adopted AI-based tools to personalise learning, automate assessment, detect plagiarism and provide tutoring at scale. This technological revolution has generated legitimate expectations regarding improvements in educational quality, the reduction of access gaps and the enhancement of institutional efficiency (Baker & Inventado, 2014; Holmes et al., 2022).

However, the very capacity of AI to generate high-quality content, automate complex tasks and simulate human behaviour has been exploited to commit next-generation academic offences. Automated plagiarism via large language models, identity theft through deepfake technology in remote examinations, the forgery of academic credentials using intelligent editing tools, and the manipulation of automated grading platforms constitute criminal modalities that overwhelm regulatory frameworks designed in a pre-digital era (Stokel-Walker, 2023; Perkins et al., 2024; Cotton et al., 2024).

In response to this landscape, the most significant normative development at the global level has been the adoption of Regulation (EU) 2024/1689 on Artificial Intelligence — the AI Act — the world's first comprehensive legally binding regulatory framework for AI, with direct and far-reaching implications for higher education institutions in their dual role as users and deployers of AI systems. However, the AI Act is not a criminal law instrument: it does not define crimes, but rather establishes obligations of transparency, risk assessment and human oversight whose violation triggers administrative liability. The articulation between this regulatory framework and national criminal law therefore remains an area of considerable legal uncertainty (Ramos, 2023; European Commission, 2024).

From an institutional perspective, the response of universities to the challenge of AI-facilitated offences has been uneven and frequently reactive. The absence of specific academic compliance policies on AI, insufficient faculty training in digital integrity, and limited coordination between the legal, technological and academic organs of institutions together configure a landscape of vulnerability that demands urgent attention (Guerrero-Doldán, 2024; UNESCO, 2023).

This article aims to systematise the landscape of AI-facilitated criminal offences in the Spanish and European university context, analyse in depth the implications of the AI Act for educational institutions, and propose a model of institutional response that integrates prevention, detection, sanction and a culture of integrity. To this end, an interdisciplinary perspective is adopted that integrates criminal law, criminology, applied ethics and educational technology studies.

2. Objectives

The objectives of this study are as follows:

1. To analyse the main typologies of criminal offences facilitated by the misuse of AI tools in university education and their accommodation within current criminal law.
2. To examine in depth Regulation (EU) 2024/1689 on AI (AI Act) and its specific obligations for higher education institutions.
3. To identify the gaps between the current regulatory framework and the real needs of institutional response to AI-facilitated offences.
4. To propose an institutional response model that integrates prevention, detection, sanction and digital integrity training.

3. Methodology

3.1. Type of study and design

This study is framed within qualitative research of an exploratory and descriptive nature, aimed at analysing the criminal offences facilitated by AI in the university setting and understanding the associated legal and ethical implications. A non-experimental, cross-sectional design was employed, combining three data collection techniques to guarantee methodological triangulation: systematic literature review, in-depth documentary analysis of the AI Act, and semi-structured interviews with experts in criminal law, criminology and educational technology (Creswell & Poth, 2018; Flick, 2021).

3.2. Population and sample

The study population includes: (1) scientific literature published in Scopus and Web of Science between 2018 and 2025; (2) the full text of Regulation (EU) 2024/1689 and its technical development documents; (3) applicable Spanish and international criminal legislation; (4) documented cases of AI-facilitated academic offences; and (5) five experts selected through purposive sampling based on demonstrated experience in technological criminal law, digital ethics and academic integrity management.

3.3. Data analysis techniques

Qualitative content analysis was employed, coding the data into thematic categories related to offence typologies, AI Act implications, prevention strategies and institutional response models. For the analysis of the AI Act, an article-by-article normative mapping of obligations relevant to the educational sector was conducted. Interviews were transcribed and subjected to thematic analysis following the procedure of Braun and Clarke (2021). Validity was guaranteed through triangulation of sources, researchers and methods.

3.4. Ethical considerations

The study respected the ethical principles of scientific research, including confidentiality and anonymity of the interviewed experts, and avoided the disclosure of data that could identify individuals involved in the cases analysed. The study design was submitted to academic integrity review prior to its execution.

4. Theoretical framework: criminal offences, artificial intelligence and university education

4.1. Concept and classification of AI-facilitated academic criminal offences

Criminal offences in the academic sphere comprise conduct that transgresses the legal and ethical norms of the educational context, undermining the integrity of the system and generating legal liability for perpetrators. The Spanish Criminal Code (Organic Law 10/1995) defines offences as deliberate or negligent actions or omissions punishable by law. In the university environment, this encompasses fraud in obtaining qualifications, forgery of academic documents, and identity theft in assessments — conduct that violates both legality and the fundamental principles of education (Giner, 2023).

The emergence of generative AI has substantially broadened the scope of these behaviours, introducing three qualitatively new elements: scale (the possibility of committing the offence massively), sophistication (production of content or documents of high quality that are barely distinguishable from legitimate ones), and accessibility (low-cost tools available to any user). The classification of AI-facilitated academic offences encompasses the following categories:

- Automated plagiarism via large language models (LLMs): generating academic work with AI and presenting it as one's own (Perkins et al., 2024).
- Identity theft via deepfakes: use of synthetic avatars or pre-recorded images to pass identity verification systems in remote examinations (ENISA, 2024).
- Forgery of academic documents: creation of false degrees, certificates

or transcripts using image generation tools (Morales et al., 2024).

- Fraud in automated assessment: manipulation of automatic grading platforms to alter results (Cotton et al., 2024).
- Production of illegal content: generation of non-consensual intimate images, scientific disinformation or hate speech (Europol, 2023).
- AI-facilitated cyberbullying: automated production of hostile messages, personalised threats or smear campaigns (Livingstone & Stoilova, 2023).

4.2. Legitimate applications of AI in higher education

To properly contextualise the risks, it is essential to acknowledge the broad spectrum of legitimate and beneficial AI applications in higher education: learning personalisation via adaptive systems, personalised intelligent tutoring, automation of academic administration, predictive analytics to identify at-risk students, accessibility and inclusion tools, and assessment systems with immediate feedback (Holmes et al., 2022; Zawacki-Richter et al., 2019; UNESCO, 2023). This dual dimension of AI — as a valuable pedagogical tool and as a potential instrument of fraud — underscores the need for regulatory frameworks that allow its benefits to be harnessed without sacrificing academic integrity.

4.3. Applicable Spanish criminal law framework

The Spanish Criminal Code contemplates a set of offence types applicable to AI-facilitated academic misconduct, albeit without explicit reference to these technologies:

- Document forgery (arts. 390–399 CC): falsification of public, private or official documents, including academic qualifications and certificates, with penalties of up to six years' imprisonment.
- Fraud (art. 248 CC): obtaining unlawful financial benefit through deception, applicable to the fraudulent acquisition of qualifications or academic funding.
- Usurpation of civil status (art. 401 CC): identity theft in assessments.
- Computer offences (arts. 197 bis, 264 CC): unlawful access to information systems and computer damage, applicable to the manipulation of assessment platforms.
- Offences against privacy (arts. 197, 197 octies CC): production and dissemination of non-consensual intimate images, as reformed by Organic Law 10/2022.
- Hate crimes (art. 510 CC): incitement to discrimination or violence via AI-generated content.

Applying these offence types to AI-facilitated misconduct frequently requires an exercise of analogical interpretation that may generate legal uncertainty, reinforcing the need for specific legislative reform that explicitly incorporates the use of AI as an aggravating circumstance or constituent element of certain offences (Ramos, 2023; Gómez-Jara Díez, 2022).

5. Regulation (EU) 2024/1689 on Artificial Intelligence (AI Act) and higher education

5.1. Context and scope of the AI Act

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on Artificial Intelligence — hereinafter the AI Act — constitutes the first comprehensive regulatory framework for AI adopted worldwide with binding legal force. Published in the Official Journal of the European Union on 12 July 2024, it entered into force on 1 August 2024, with a phased application period extending to 2 August 2027 for the most

complex provisions. The AI Act applies territorially throughout the European Economic Area and has extraterritorial effect on AI systems marketed or used in the EU regardless of the provider's location, which is particularly relevant for academic institutions using AI platforms developed outside Europe (European Commission, 2024).

The AI Act is not a criminal law instrument: it does not criminalise conduct nor establish criminal penalties. Its objective is to prevent risks associated with AI through obligations of transparency, conformity assessment, human oversight and accountability, under the principle of regulatory proportionality. The administrative sanctions provided for can reach up to 35 million euros or 7% of global annual turnover for the most serious infringements (Article 99 AI Act). Nevertheless, non-compliance with AI Act obligations may have indirect implications in the criminal sphere, insofar as it evidences deficient institutional organisation relevant to the purposes of Article 31 bis of the Spanish Criminal Code (Gómez-Jara Díez, 2022).

5.2. Risk classification and its application in education

The AI Act articulates its regulatory regime around a classification of AI systems into four risk levels. This classification has direct consequences for the obligations falling on providers and deployers of AI systems in the university environment.

The most relevant category for educational institutions is that of high risk. Annex III of the AI Act expressly includes in this category AI systems used to determine access to education and vocational training, to evaluate students (including the detection of fraudulent behaviour), and for guidance during the learning process. This classification means that most remote proctoring platforms, adaptive assessment systems and AI-based university admissions tools must comply with the requirements of the high-risk regime (European Commission, 2024).

Table 1 summarises the risk levels and their implications for higher education:

Risk level	Category	Examples in education	Main obligations
Unacceptable (prohibited)	Mass social scoring; subliminal manipulation systems	Behavioural monitoring of students without legal basis	Absolute prohibition on use
High risk	Assessment a selection systems in education (Annex III)	Proctoring platforms, automated grading systems AI-based admissions	Conformity assessment, human oversight, EU database registration, transparency
Limited risk	Systems with transparency obligations	Tutorial chatbo academic virtua assistants	Inform users they are interacting wi AI
Minimal risk	Majority of educational A applications	Learning personalisation tools, spell checkers	No specific additional obligations

Table 1. AI Act risk levels and implications for higher education (adapted from European Commission, 2024).

5.3. Obligations for higher education institutions as deployers of AI

Universities that use high-risk AI systems in their academic processes act as deployers within the meaning of the AI Act (Article 3.4) and are subject to a specific set of obligations that merit detailed analysis:

5.3.1. Fundamental Rights Impact Assessment (FRIA)

Article 27 of the AI Act obliges deployers of high-risk AI systems in the public sector — a category that encompasses public universities — to carry out a Fundamental Rights Impact Assessment (FRIA) before putting the system into operation. This assessment must identify risks to rights such as non-discrimination, privacy, data protection or access to education, and the measures adopted to mitigate them. For private universities, the FRIA is recommended but not legally mandatory, although its voluntary execution

may constitute a relevant factor for civil and criminal liability purposes (Article 27.4 AI Act).

The FRIA is particularly critical in the case of proctoring systems, which involve the mass processing of biometric and behavioural data from students. Previous studies have shown that these systems exhibit significantly higher error rates for students of certain ethnicities or with specific health conditions, potentially generating situations of algorithmic discrimination with legal relevance (Nigam et al., 2021; Holstein & Doroudi, 2022).

5.3.2. Effective human oversight

Article 14 of the AI Act establishes that high-risk AI systems must be designed to allow effective human oversight during their use. For educational institutions, this means that no academic decision with significant consequences for the student — suspension, exclusion, final grade, access to selective programmes — can be taken in an exclusively automated manner. There must always be a qualified human review capable of detecting system failures, questioning its results and taking the final decision with full responsibility (European Commission, 2024).

This requirement has direct implications for the design of university assessment processes. Institutions that delegate final grading to automated systems without human review not only breach the AI Act, but may also violate students' right to effective academic protection and the principle of accountability in the processing of personal data recognised under the GDPR (Regulation (EU) 2016/679).

5.3.3. Transparency and student information

Article 50 of the AI Act establishes transparency obligations for AI systems that interact directly with people. Educational institutions using tutorial chatbots, learning assistants or automated feedback systems must clearly inform students that they are interacting with an AI system, unless this is obvious from the context. This transparency obligation extends to the use of AI-based fraud or plagiarism detection systems: students must know that their work is subjected to automated analysis and have access to a comprehensible explanation of the criteria and results (Mökander et al., 2023).

The absence of this information may violate the transparency principle of the GDPR (Article 13) and generate institutional liability towards students affected by decisions based on insufficiently explained AI systems.

5.3.4. Data governance and non-discrimination

Article 10 of the AI Act imposes on providers of high-risk AI systems — and to a certain extent on their deployers — data governance obligations that include ensuring the quality, representativeness and freedom from bias of the datasets used to train and operate the system. For universities, this implies an active responsibility to verify that the AI systems they acquire or use have been trained with data representative of their student population and do not perpetuate discriminatory biases based on gender, ethnicity, disability or other protected characteristics (Holstein & Doroudi, 2022; UNESCO, 2023).

This obligation connects directly with the principle of non-discrimination enshrined in Article 14 of the Spanish Constitution and with the right to equal opportunities in access to education. Institutions that use discriminatory AI systems without adopting corrective measures may incur civil and administrative liability, and in serious cases, criminal liability.

5.3.5. Activity logging and technical documentation

Articles 12 and 73 of the AI Act oblige deployers of high-risk AI systems to maintain records of activities performed with the system for a minimum period of six months, as well as to retain the technical documentation provided by the provider. For educational institutions, this means implementing logging systems that record the operation of automated assessment tools, plagiarism detection and proctoring, so that decisions taken with AI support can be audited

ex post and potential anomalies or misuse detected (European Commission, 2024).

This logging obligation is particularly relevant in the context of academic offences: insofar as records allow the fraudulent use of AI systems to be traced, they constitute both a preventive tool and evidence in disciplinary and criminal proceedings.

5.4. Absolute prohibitions with impact on the university environment

Article 5 of the AI Act establishes a catalogue of AI practices that are absolutely prohibited as incompatible with fundamental rights. Among those with direct relevance to the university environment:

- Generalised social scoring systems: the prohibition on AI systems that assign trustworthiness or social behaviour scores to people on the basis of accumulated data across multiple contexts is relevant for any university platform that might implement AI-based global academic reputation systems.
- Subliminal manipulation: the use of AI techniques that manipulate people's behaviour imperceptibly and to their detriment is prohibited, which may affect certain gamification applications and automated incentive systems on educational platforms.
- Emotion inference in workplace and educational contexts: Article 5.1.f prohibits the commercialisation and use of AI systems to infer the emotions of individuals in work and educational settings, except for duly justified medical or safety reasons. This prohibition directly affects emotion detection systems used in some adaptive learning environments.

5.5. The AI Act as an instrument for preventing academic offences

Although the AI Act was not specifically designed as an instrument for preventing academic fraud, its obligations have a significant preventive effect on AI-facilitated offences. The obligation of effective human oversight reduces the space for fraud in automated assessments. Transparency about the use of AI makes the impunity of technological plagiarism more difficult. Data quality and non-discrimination requirements limit the proliferation of biased assessment systems susceptible to manipulation. And activity logging provides essential documentary evidence for the investigation of offences (Mökander et al., 2023; European Commission, 2024).

However, the AI Act presents significant limitations as an instrument for preventing academic fraud. Its focus is centred on institutional providers and deployers of AI systems, not on end users (students), who are frequently the active subjects of the offences. Furthermore, the AI Act's obligations apply only to AI systems formally deployed by institutions, without covering students' private use of external AI tools — which is precisely the most widespread modality of AI-facilitated academic fraud (Perkins et al., 2024).

This structural limitation of the AI Act reinforces the need for a complementary institutional approach focused on cultural prevention and the redesign of assessment processes, which is analysed in the following section.

5.6. AI Act implementation timeline and compliance deadlines

The application of the AI Act is phased. The most relevant deadlines for educational institutions are: (1) August 2025: entry into force of the absolute prohibitions of Article 5, including the prohibition of emotion inference systems; (2) August 2026: application of obligations relating to general-purpose AI models (GPAI), relevant for institutions that develop or adapt language models for educational use; and (3) August 2027: full application of obligations for high-risk systems, including FRIA obligations, human oversight, transparency and logging that affect assessment and proctoring systems

(European Commission, 2024). Educational institutions must begin their adaptation process immediately, starting with an inventory of AI systems in use and a preliminary risk assessment in accordance with the AI Act's taxonomy.

6. Documented criminal offences and trend analysis

6.1. Documented cases

Through the analysis of academic, legal and journalistic sources, multiple cases were identified in which AI facilitated the commission of criminal offences in the university setting:

1. Automated plagiarism with LLMs: Cases documented at Spanish, British and American universities in which students used GPT-4 and similar models to produce undergraduate dissertations and research articles. Instructors detected stylistic inconsistencies and fabricated references (model hallucinations), leading to disciplinary investigations and, in some cases, criminal complaints for academic fraud (Perkins et al., 2024).
2. Identity theft via deepfake in remote examinations: European and North American universities reported incidents in which AI avatars or pre-recorded images passed identity verification systems on proctoring platforms. Some cases involved the use of silicone masks and real-time filters to deceive facial recognition systems (ENISA, 2024; Nigam et al., 2021).
3. Forgery of academic documents: Creation of university diplomas and academic transcripts using image diffusion models (Stable Diffusion, DALL-E), detected by employers when verifying credential authenticity (Morales et al., 2024).
4. Non-consensual intimate images in the school environment: In 2025, the arrest was recorded in Palma (Spain) of a minor who used AI to generate intimate images of female classmates from photographs on their social media profiles. The case was classified as an offence against sexual integrity (arts. 197 octies and 189 CC, as reformed by Organic Law 10/2022) and referred to the Juvenile Prosecution Service.
5. Fraud on automated assessment platforms: Students exploited vulnerabilities in adaptive assessment platforms, intercepting communications between client and server to submit pre-defined answers without passing through the legitimate assessment module (Cotton et al., 2024).
6. Fabrication of research data: Use of AI to generate bibliographic references, experimental data and non-existent research results in doctoral theses and articles submitted to scientific journals. This conduct violates the principles of scientific integrity and may constitute document forgery when it affects research funded by public funds (Gao et al., 2024).

6.2. Trend analysis

The data analysed allow the identification of five main trends: (1) increasing technological sophistication of offences, making detection with conventional tools progressively more difficult; (2) diversification of the criminal spectrum, extending well beyond traditional textual plagiarism; (3) sustained growth in reported cases between 2022 and 2025, particularly in online education environments; (4) growing difficulties in regulation and detection arising from the rapid evolution of AI models; and (5) emerging institutional awareness, although with significant inequalities between institutions and countries in terms of resources and capacities (ENISA, 2024; Guerrero-Doldán, 2024).

7. Prevention and institutional response to AI-facilitated offences

7.1. Principles of an institutional response model

An effective institutional response to AI-facilitated academic offences cannot be limited to reactive measures of detection and sanction. It must be articulated as an integrated system operating at four complementary levels: prevention (reducing the probability of the offence being committed), detection (identifying offences that have occurred), sanction (responding proportionately and deterrently) and institutional learning (continuously improving the system based on detected cases). This systemic approach is the one that has demonstrated greatest efficacy in the corporate compliance field and must be transposed to the academic context (Bretag et al., 2019; Gómez-Jara Díez, 2022).

7.2. Prevention: training, policy and assessment redesign

7.2.1. Training in digital integrity and AI ethics

Training in digital ethics, academic integrity and responsible AI use is the preventive measure with the greatest long-term impact, as consistently highlighted by both the specialist literature and the experts interviewed. This training must target all members of the university community: students (from the moment of enrolment), academic staff, and administrative and support personnel. Training programmes must be specific to the AI context — not limited to traditional academic integrity approaches — and must incorporate practical case studies, ethical reflection on the implications of fraud, and knowledge of the legal consequences of the misuse of AI tools (McCabe et al., 2012; ICAI, 2021; UNESCO, 2023).

7.2.2. Updating institutional policies

Universities must update their disciplinary regulations to explicitly incorporate AI-facilitated offences as typified conduct, with clear investigation procedures, resolution timelines and proportionate sanctions. These policies must establish with clarity which uses of AI are permitted for each type of academic task, distinguishing between assisted use (permitted with a declaration of use), supervised use (subject to specific conditions) and prohibited use (submission of AI-generated content as one's own without attribution). Normative ambiguity is one of the main factors that encourages academic fraud (Bretag et al., 2019; Guerrero-Doldán, 2024).

7.2.3. Redesigning assessment tasks

Redesigning assessment tasks to make them resistant to technological fraud is the most effective prevention strategy from a technical perspective. Assessments should incorporate elements that cannot plausibly be produced by an AI system: references to the student's own experience, analysis of unpublished materials not available in public databases, reflection on recent debates in the field, oral defence or real-time justification of the elaboration process. Continuous assessment, based on multiple learning evidences distributed throughout the semester, is significantly more resistant to fraud than single examinations or final projects (Cotton et al., 2024; Perkins et al., 2024).

7.3. Detection: tools and limitations

AI-generated content detection systems have developed rapidly, but present significant limitations that institutions must be aware of. The most widely used tools (GPTZero, Turnitin AI Detection, Copyleaks) show relevant error rates: the most rigorous studies reveal false positives — human texts incorrectly classified as AI-generated — that may harm innocent students, and false negatives — AI-generated texts that evade detection — that allow fraud to continue (Weber-Wulff et al., 2023). These limitations mean that the results of AI detectors should not be used as sole evidence in disciplinary proceedings, but rather as an indicator that activates a broader

investigation based on multiple forms of evidence.

Institutions must also implement systems for logging and auditing the use of AI platforms in their institutional environments, as required by the AI Act for high-risk systems. These records can provide relevant documentary evidence in the investigation of offences and constitute a deterrent measure in themselves.

7.4. Sanction and academic compliance

Sanctions must be applied consistently, proportionate to the seriousness of the infringement and with full respect for the principle of due process. The sanctioning regime must include a gradation of consequences based on intentionality, recidivism and the impact of the offence on third parties. The most severe sanctions — including expulsion and the annulment of fraudulently obtained qualifications, with the consequent criminal complaint — must be reserved for the most serious and deliberate cases (Bretag et al., 2019).

The concept of academic compliance — transferred from the corporate sphere to the educational context — implies that institutions implement a comprehensive programme of prevention, detection and response that complies with AI Act obligations and the Criminal Code. This programme must include an anonymous reporting channel, an academic integrity officer with specific AI competencies, periodic reviews of the efficacy of measures adopted, and an annual transparency report on offences detected and sanctions imposed (Gómez-Jara Díez, 2022).

7.5. Proposals for legislative reform

Beyond institutional responses, the effective combating of AI-facilitated academic offences requires specific legislative updates:

- Incorporating the use of AI as an aggravating circumstance in offences of document forgery, fraud and identity theft when such use significantly increases the capacity for harm or the difficulty of detection.
- Developing specific regulations governing the use of AI tools in academic assessments, establishing the conditions for legitimate use and the consequences of misuse.
- Establishing obligations to notify competent authorities of serious AI-facilitated academic fraud incidents, analogous to the data breach notification obligations of the GDPR.
- Creating an anonymised national registry of AI-facilitated academic fraud incidents, enabling analysis of trends and evaluation of the efficacy of preventive measures at the national level.
- Promoting within the Council of Europe the development of an international Convention on academic integrity in the AI era, providing a common framework for cross-border cooperation in prosecuting these offences.

8. Discussion

8.1. Gaps between the regulatory framework and institutional reality

The research findings confirm the existence of a significant gap between the formally applicable regulatory framework — Criminal Code, AI Act, GDPR — and the real capacity of educational institutions to prevent, detect and sanction AI-facilitated offences. This gap operates across three dimensions: a knowledge gap (many institutional managers are unaware of their obligations under the AI Act or the extent of corporate criminal liability), a resources gap (institutions frequently lack the technical and human capacities to implement the detection and logging systems required), and a cultural gap (digital ethics and academic integrity in the AI context are not sufficiently integrated into institutional culture) (Guerrero-Doldán, 2024; ICAI, 2021).

8.2. The AI Act as an opportunity and a challenge

The AI Act should be valued not only as a set of compliance obligations, but as an opportunity for educational institutions to rethink their relationship with AI in a systematic and values-driven manner. The impact assessment, human oversight and transparency obligations imposed by the AI Act are, ultimately, expressions of ethical principles that should inspire any responsible use of AI in education: respect for student autonomy, guaranteeing fairness in assessment, and accountability for decisions that affect people's academic and professional futures. The implementation of the AI Act can thus become a catalyst for a broader institutional transformation towards a culture of digital integrity (Mökander et al., 2023).

At the same time, the AI Act presents significant challenges for smaller educational institutions with more limited resources: the cost of conformity assessments, logging systems and technical documentation may represent a disproportionate burden for small universities or those with limited technological management capacity. This asymmetry requires active intervention by the ministries responsible for education and digitalisation, which must provide sector-specific implementation guidelines, technical support mechanisms and, where appropriate, proportionate adaptations of the obligation regime (European Commission, 2024).

8.3. Limitations of the study

- Data availability: many offences are not reported for reputational reasons or due to the absence of registration protocols.
- Normative variability: legislation varies significantly between countries, making direct comparisons difficult.
- Qualitative approach: does not allow the precise quantification of the global incidence of offences or objective measurement of the impact of preventive measures.
- Rate of change: the rapid evolution of AI models may render some statements outdated in the short term.

These limitations point to the need for complementary quantitative and longitudinal studies, tracking cohorts of institutions through the AI Act adaptation process.

9. Conclusions

1. AI has profoundly transformed the university environment, introducing both significant pedagogical opportunities and new modalities of criminal offences that overwhelm traditional regulatory frameworks. The challenge is not essentially technological, but institutional, ethical and legal: it requires the construction of a robust culture of digital integrity, sustained by updated regulatory frameworks and adequate institutional capacities.
2. The main typologies of AI-facilitated academic offences — automated plagiarism, identity theft, document forgery, fraud in automated assessment and production of illegal content — can be accommodated within the current Spanish Criminal Code, albeit with significant interpretive difficulties that generate legal uncertainty and risk of impunity. Specific legislative reform incorporating the use of AI as an aggravating circumstance is necessary and urgent.
3. The AI Act establishes a reference regulatory framework with direct and relevant obligations for higher education institutions: fundamental rights impact assessment, effective human oversight of assessment systems, student transparency, data governance and activity logging. Compliance with these obligations is, moreover, an opportunity to

systematically transform institutions' relationship with AI.

4. However, the AI Act presents structural limitations as an instrument for preventing academic fraud: it is centred on institutional deployers of AI, not on students' private use of external tools. This reinforces the need for complementary strategies focused on ethical training, assessment redesign and academic compliance.
5. The most effective institutional response integrates four levels: prevention (training, policy and assessment redesign), detection (technological tools with awareness of their limitations, audit records), sanction (proportionate and transparent disciplinary procedures) and institutional learning (continuous improvement based on detected cases). This academic compliance model should draw on best practices from corporate compliance, adapted to the specific context of higher education.
6. Future research must advance towards quantitative and longitudinal studies that measure the real incidence of offences, evaluate the comparative efficacy of prevention strategies and analyse the real impact of AI Act implementation on the transformation of university institutional practices.

Academic integrity and ethical responsibility in the use of AI must be fundamental pillars of the twenty-first century university. Only thus can it be guaranteed that the adoption of these technologies contributes to students' academic and professional development without compromising the legality, equity and values that give meaning to higher education.

Referencias

BAKER, Ryan. S., & INVENTADO, Paul. Salvador. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), Learning analytics: From research to practice (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4 PMCid:PMC4355053

BRAUN, Virginia., & CLARKE, Victoria. (2021). Thematic analysis: A practical guide. SAGE Publications.

BRETAG, Tracey., HARPER, Rowena., BURTON, Michael., ELLIS, Cath., NEWTON, Philip., ROZENBERG, Pearl., SADDIQUI, Sonia., & VAN HAERINGEN, Karen. (2019). Contract cheating and assessment design: Exploring the relationship. Assessment & Evaluation in Higher Education, 44(5), 676–691. https://doi.org/10.1080/02602938.2018.1527892

COTTON, Debby. R. E., COTTON, Peter. A., & SHIPWAY, J. Reuben. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

CRESWELL, John. W., & POTH, Cheryl. N. (2018). Qualitative inquiry and research design: Choosing among five approaches (4th ed.). SAGE Publications.

DELGADO MORÁN, Juan. José. y GINER ALEGRÍA, Cesar. Augusto. (2017). La protección de datos de carácter personal en el ordenamiento jurídico español. Sociedad, empresas e instituciones: una aproximación. Pp. 347-365. Sotec Editorial

DELGADO MORÁN, Juan. José. (2024). Acoso y agresión en las nuevas tecnologías: ciberacoso/ciberodio. *AlmaMater. Cuadernos de Psicosociobiología de la Violencia: Educación y Prevención*, nº 5, Dykinson, pp. 107-122. https://doi.org/10.14679/3315

ENISA. (2024). Artificial intelligence cybersecurity challenges: Threat landscape for artificial intelligence. European Union Agency for Cybersecurity. https://www.enisa.europa.eu/publications/e

nisa-threat-landscape-for-artificial-intelligence

European Commission. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689

Europol. (2023). Facing reality? Law enforcement and the challenge of deepfakes. Europol Innovation Lab. https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes

FLICK, Uwe. (2021). Introducing research methodology: Thinking your way through your research project (3rd ed.). SAGE Publications.

GINER ALEGRÍA, Cesar. Augusto., & DELGADO MORAN, Juan. José. (2017). Consideraciones criminológicas sobre el perfil del stalker y el acecho mediante ciberstalking. *Estudios en seguridad y defensa*, 12(24), 19-35. https://doi.org/10.25062/1900-8325.250

GINER ALEGRÍA, Cesar. Augusto. (2023). La responsabilidad penal en el uso de inteligencia artificial: Análisis desde el derecho español. *Cuadernos de Política Criminal,* 140, 87–112.

GINER ALEGRÍA, César Augusto, & GÓMEZ ABRIL, Francisco. (2026). Inteligencia artificial y su relación actual con las ciencias de la seguridad: De la ficción a la realidad. *Quadernos de criminología: revista de criminología y ciencias forenses,* Nº. 67, págs. 27-39

GINER ALEGRÍA, César Augusto. (2026). Ilícitos penales en enseñanzas universitarias al amparo de la IA. *Revista de educación y derecho*, (33). https://doi.org/10.1344/REYD2026.33.54035

GINER ALEGRÍA, César Augusto. (2025). Neurociencias y derecho penal. Diario La Ley, Nº 10652,

GÓMEZ-JARA DÍEZ, Carlos. (2022). Corporate criminal liability and compliance programs in Spain. In M. Pieth & R. Ivory (Eds.), Corporate criminal liability: Emergence, convergence, and risk (pp. 265–292). Springer. https://doi.org/10.1007/978-90-481-3308-0_12

HOLMES Wayne, Kaska PORAYSKA-POMSTA, Ken HOLSTEIN, Emma SUTHERLAND, Toby BAKER, Simon BUCKINGHAM SHUM, Olga C. SANTOS, Mercedes T. RODRIGO, Mutlu CUKUROVA, Ig Ibert BITTENCOURT, & Kenneth R. KOEDINGER. (2022). Ethics of AI in education: Towards a community-wide agenda. International Journal of Artificial Intelligence in Education, 32(4), 504–526. https://doi.org/10.1007/s40593-021-00239-1

HOLSTEIN, Kenneth., & DOROUDI, Shayan. (2022). Equity and artificial intelligence in education: Will 'AIEd' amplify or alleviate inequities in education? In W. Holmes & K. Porayska-Pomsta (Eds.), The ethics of artificial intelligence in education (pp. 151–179). Routledge. https://doi.org/10.4324/9780429329067-10

International Center for Academic Integrity [ICAI]. (2021). The fundamental values of academic integrity (3rd ed.). Clemson University. https://academicintegrity.org/images/pdfs/20019_ICAI-Fundamental-Values_R12.pdf

LIM, Weng. Mark., GUNASEKARA, Asanka., PALLANT, Jessica. Leigh., katerina, Jason. Ian., & PECHENKINA, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. The International Journal of Management Education, 21(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

LIVINGSTONE, Sonia., & STOILOVA, Mariya. (2023). Children and digital technologies: Risks and opportunities in AI-driven environments. Oxford University Press.

LIZ RIVAS, Lenny. (2024).Violencia y agresión entre iguales a través de las TICS:

Cyberbulling. AlmaMater. Cuadernos de Psicosociobiología de la Violencia: Educación y Prevención, nº 5, 2024, Dykinson, pp. 89-105. https://doi.org/10.14679/3314

MAZURIER, Pablo, A., DELGADO MORÁN, Juan, José & PAYA SANTOS, Claudio, A. (2019). Gobernanza constructivista de la internet. *Teoría y Praxis*, 17(34), 107-130. https://doi.org/10.5377/typ.v1i34.14823

MCCABE, Donald. L., BUTTERFIELD, Kenneth. D., & TREVIÑO, Linda. K. (2012). Cheating in college: Why students do it and what educators can do about it. Johns Hopkins University Press.

MÖKANDER, Jakob., & FLORIDI, Luciano. (2023). Operationalising AI governance through ethics-based auditing: An industry case study. AI and Society, 38(2), 539–550. https://doi.org/10.2139/ssrn.4268361

MORALES TIRADO, Alba., MULHOLLAND, Paul., & FERNÁNDEZ, Miriam. (2024). Towards an operational responsible AI framework for learning analytics in higher education. arXiv. https://arxiv.org/abs/2410.05827

NIGAM, Aditya., PASRICHA, Rhitvik., SINGH, Tarishi., & CHURI, Prathamesh. (2021). A systematic review on AI-based proctoring systems: Past, present and future. Education and Information Technologies, 26(5), 6421–6445. https://doi.org/10.1007/s10639-021-10597-x PMid:34177348 PMCid:PMC8220875

PAYÁ SANTOS, Claudio. Augusto, & DELGADO MORÁN, Juan. José (2025). Cognitive biases: Understanding and mitigating their effects. American Based Research Journal. Vol (14). Issue, 4 https://doi.org/10.5281/zenodo.15412164

PAYÁ SANTOS, Claudio. Augusto, & DELGADO MORÁN, Juan. José;. MARTINO, Luigi; GARCÍA SEGURA, Luis, A.; DIZ CASAL, Javier, & FERNÁNDEZ-RODRÍGUEZ, Juan, Carlos. (2023). Fuzzy Logic analysis for managing Uncertain Situations. Review of Contemporary Philosophy Vol 22 (1), 2023 pp. 6780 -6797. https://doi.org/10.52783/rcp.1132

PERKINS, Mike., ROE, Jasper., POSTMA, Darius., MCGAUGHRAN, James., & HICKERSON, Don. (2024). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. Journal of Academic Ethics, 22(1), 89–113. https://doi.org/10.1007/s10805-023-09492-6

SANZ GONZÁLEZ, Roger, LUQUE JUÁREZ, José M.ª, MARTINO, Luigi, LIZ RIVAS, Lenny, DELGADO MORÁN, Juan José & PAYÁ SANTOS, Claudio Augusto. (2024) Artificial Intelligence Applications for Criminology and Police Sciences. *International Journal of Humanities and Social Science.* Vol. 14, No. 2, pp. 139-148. https://doi.org/10.15640/jehd.v14n2a14

STOKEL-WALKER, Chris. (2023). AI bots like ChatGPT could be a problem for universities — but the sky isn't falling. Nature, 613, 620–621. https://doi.org/10.1038/d41586-023-00107-z PMid:36653617

UNESCO. (2023). Guidance for generative AI in education and research. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

WEBER-WULFF, Debora., ANOHINA-NAUMECA, Alla., BJELOBABA, Sonja., Foltýnek, Tomás., GUERRERO-DOLDÁN, Jean., POPOOLA, Olumide, ŠIGUT, Petr & WADDINGTON, Lorna. (2023). Testing of detection tools for AI-generated text. International Journal for Educational Integrity, 19(1), 26. https://doi.org/10.1007/s40979-023-00146-z

ZAWACKI-RICHTER, Olaf., MARÍN, Victoria. I., BOND, Melissa., & GOUVERNEUR, Franziska. (2019). Systematic review of research on artificial intelligence applications in higher education — where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0