From Recommender Systems to Anti-Fraud: Practices for Implementing AI in Financial and Educational Products, Taking into Account Regulatory Restrictions

Daria Bogun
Russia, Moscow
CEO/Founder of LearnHub

Abstract— *The article examines the practical experience of implementing artificial intelligence in financial and educational products, taking into account regulatory constraints. The relevance of the study is driven by the rapid expansion of AI applications in lending, payments, certification, and the creation of adaptive learning trajectories, as well as the necessity of strict compliance with the AI Act, PSD2, DORA, GDPR, FERPA, and COPPA, all designed to protect citizens' economic and educational rights. The work aims to perform a comparative analysis of legal frameworks and to identify effective technical and organizational practices for deploying recommendation systems, anti-fraud frameworks, and adaptive learning solutions under stringent regulation. The novelty of the study lies in the comprehensive comparison of European and American directives regarding the classification of high-risk scenarios, mandatory algorithm explainability, and data protection, as well as the synthesis of best practices from the corporate cases of Visa, Mastercard, and Trinetix, alongside publications by the World Bank and Turnitin. A unified technical architecture shall include a feature store with logs, an explainable AI layer, and an ML-Ops framework controlled by compliance. Three major vectors come out: utmost precision in classifying the risk, transparency on how decisions are made, and strictly minimized processed data. Practically, this means pseudonymization, federated learning, differential privacy, synthetic data generation, and model quality and drift monitoring that will be compliant with regulations, as well as ensuring the trust of regulators and users. This article will be helpful for developers, compliance specialists, product managers, and researchers in the legal regulation and AI implementation within the financial and educational sectors.*

Keywords— *Artificial intelligence, financial products, educational products, regulation.*

I. Introduction

Artificial intelligence has ceased to be an experimental tool and has become an infrastructural layer for services managing both money and knowledge: according to a global McKinsey survey, 78% of organizations now use AI in at least one business function, compared with 72% a year earlier [1]. At the same time, financial platforms and educational institutions bear a unique burden, since processes such as lending, payments, certification, and knowledge assessment directly affect citizens' economic rights and demand especially stringent regulatory protection.

In the banking and insurance sectors, AI performs several critically essential tasks: it recommends products to clients, calculates risk profiles in milliseconds, and, when necessary, blocks suspicious transactions until KYC verification is completed. Not surprisingly, in a Morgan Stanley survey, the proportion of firms using generative models rose from 48% to 71% among insurers and from 66% to 73% among financial services between January and July 2025, confirming AI's shift from a sandbox to productive anti-fraud frameworks and targeted-offer systems [2].

Education is undergoing a similar transformation, yet from a different perspective: here, AI accelerates the creation of instructional materials, generates adaptive learning trajectories, and alerts to potential academic risks. An IDC study, cited in a special Microsoft report, shows that 86% of educational organizations already employ generative models, a record across sectors [3]. Moreover, six out of ten educators in the United States use AI tools in their daily work, and one in three does so weekly, saving on average almost six hours per week [4].

II. Materials and Methodology

This study is based on the analysis of 20 key sources, encompassing global surveys, industry reports, technology company case studies, academic publications, and regulatory acts. As its empirical foundation, it uses the McKinsey global survey showing that 78% of organizations deploy AI in business functions [1]; Morgan Stanley data on generative model usage growth in insurance and financial services from 48% to 71% and from 66% to 73%, respectively, between January and July 2025 [2]; the Microsoft report recording a record 86% AI adoption in educational institutions [3]; and the Gallup survey indicating that three in ten educators save nearly six hours weekly via AI tools [4]. Corporate cases from Visa and Mastercard demonstrated the effectiveness of anti-fraud systems based on generative and graph models [14, 15]. At the same time, analytics from Trinetix and McKinsey revealed the commercial potential of generative AI in banking scoring and personalization [13, 16]. Reports from the World Bank [17], K-12 Dive and Education Week [18, 19] enabled assessment of educational risks and the rise of deep-fake content [20]. The regulatory components are comprised of the AI Act [5], PSD2 [6], DORA [7], AMLD VI [8], the GDPR Enforcement Tracker

[9], FERPA [10], COPPA [11], in addition to the U.S. Department of Education guidelines dated July 22, 2025 [12].

Comparative analysis of the EU AI Act, and financial directives (PSD2, DORA, AMLD VI) to GDPR, FERPA, and COPPA is applied; content analysis of industry and corporate reports from Visa, Mastercard, and Trinetix sheds further light on feature-store usage, feature logging practices as well as Explainable AI practices [13-15]; quantitative analysis of survey data from McKinsey, Morgan Stanley, Microsoft, Gallup and K-12 Dive all concerning scales of adoption of AI as well as its economic impact [1-4, 18]; desk based research drawing on academic and industry publications by World Bank as well as Turnitin regarding issues covering data protection pseudonymization methods deep-fake risks machine-generated text detection [17, 19, 20]; synthesis technical security measures— federated learning differential privacy synthetic dataset generation—in minimizing risks ensuring transparency supporting model auditing.

III. Results and Discussion

Regulatory requirements for artificial intelligence form two clusters that are closely linked by risk-management logic, yet derive from different legal sources. In Europe, the AI Act provides the unified foundation: it defines high-risk scenarios—such as credit scoring and automated student assessment—and mandates model registration, conformity assessment, and publication of detailed documentation on data and metrics [5]. For financial organizations, these requirements overlay the PSD2 payment directive, which instituted strong customer authentication and, according to the EBA, reduced credential-based fraud by 20% in its first two years of application [6]. Effective January 17, 2025, DORA obliges banks and investment firms to continuously test IT resilience, include AI models in disaster-recovery plans, and retain event logs for five years [7]. The Sixth Anti-Money Laundering Directive expands the list of entities subject to customer due diligence and explicitly permits algorithmic transaction monitoring as a due diligence tool, provided regular bias audits are conducted [8].

In education, the baseline remains GDPR: according to the CMS Enforcement Tracker, the Public Sector and Education segment accounts for over €30 million in fines, underscoring the seriousness of student privacy issues, as shown in Figure 1 [9].

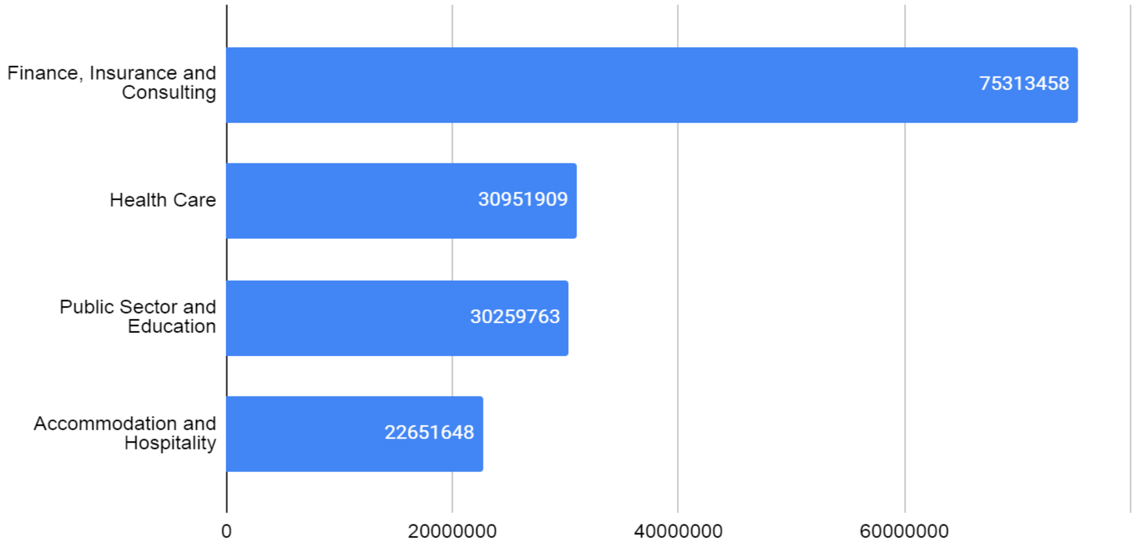

Fig. 1. Sectoral Distribution of Cumulative Fines: Total Monetary Value and Frequency of Imposed Penalties [9]

In the United States, student rights are protected by FERPA, which classifies educational data as educational records and restricts their disclosure to third parties without written consent [10]. For audiences under 13 years of age, COPPA applies: AI-platform operators must obtain verifiable parental consent before collecting any personal information, and the regulator has proposed tightening the rules by disabling targeted advertising by default and limiting push notifications that encourage prolonged service use [11]. An additional layer is provided by the U.S. Department of Education's federal guidance of July 22, 2025, which permits funding for AI solutions only if the algorithm's logic is described transparently and parents are allowed to review the datasets used [12].

Comparing both domains reveals three common vectors. First, risk classification: the AI Act assigns financial scoring and knowledge assessment to a single high-risk category, whereas in the United States, similar outcomes are achieved through the combination of FINRA, COPPA, and FERPA, effectively establishing the same entry barriers. Second, explainability: in finance, this is a stringent requirement under DORA and FINRA; in education, it is a mandatory notification under FERPA and the recent Department of Education recommendations. Third, data protection: PSD2 and AMLD VI focus on the integrity and non-disclosure of payment attributes, whereas educational regulations place primary importance on the user's age and the presence of parental controls. Thus, despite differing sectoral emphases, practical compliance implementation converges on a unified technical framework: minimizing input data, conducting regular bias audits, retaining logs for post-hoc explanations, and being prepared to suspend a model if its metrics drift beyond acceptable limits.

The growth of AI use in financial products began with personalized recommendations, where ranking models integrate account history, real-time events, and signals from third-party sources to propose the next best action in a mobile banking app in a matter of seconds. To the customer, this appears as neutral advice on a savings account or micro-insurance, yet internally, feature-store pipelines transform data to comply with PSD2 requirements and log all features for subsequent compliance review. The economic impact is measurable: McKinsey estimates that personalization and generative AI could add up to USD 340 billion in annual revenue to banks, roughly 4.7% of industry revenues [13]. Because the AI Act classifies financial-product recommendation systems as high-risk, developers must construct calculation chains—derivative → feature contribution → final score—in a format suitable for automatic export to compliance registers.

The following layer concerns transaction monitoring and anti-fraud. Generative models analyze payment-network topology, the distance between the customer node and the point-of-sale, device behavior, and even CVV entry speed, to decide authorization. During the Christmas peak of 2024, Visa's architecture of this type declined 134 million suspicious transactions and blocked 85% more fraudulent attempts than the previous year, as shown in Figure 2 [14].

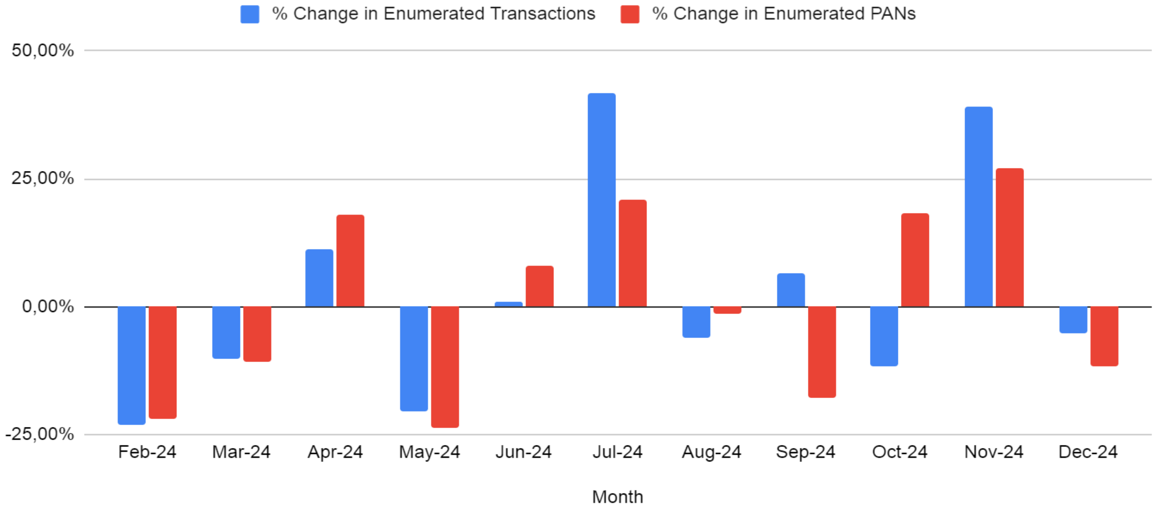

Fig. 2. Temporal Dynamics of Fraudulent Transaction and PAN Enumeration Rates [14]

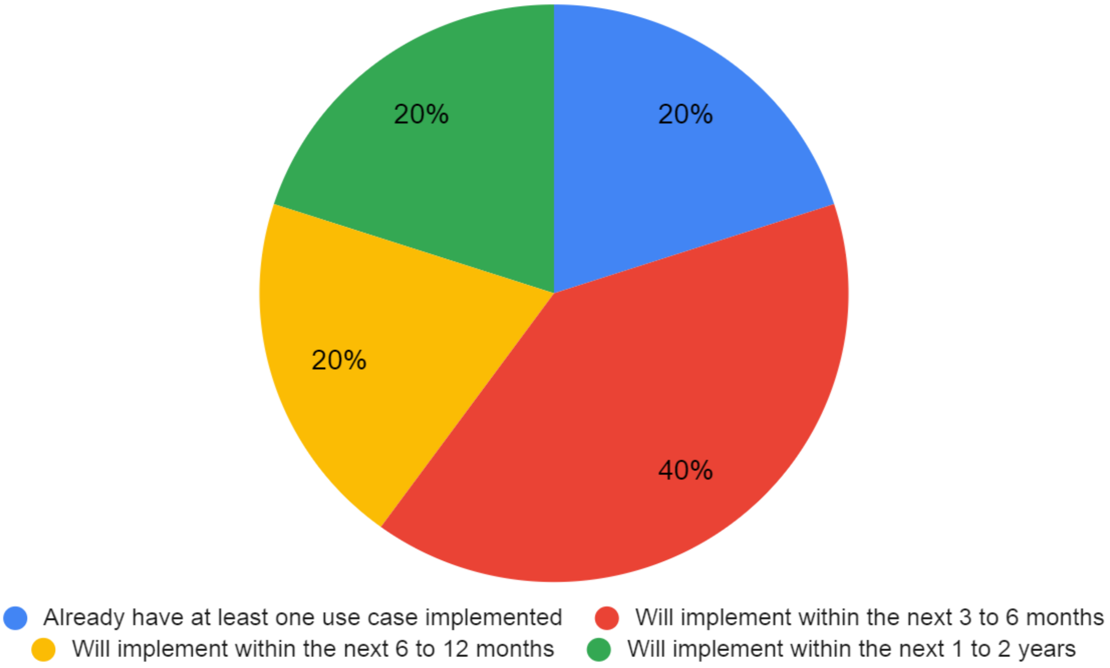

Fig. 3. Use-Case Deployment Timeline Among Respondents [16]

Mastercard adopted graph-based representations: its Decision Intelligence Pro scans a trillion parameters in 50 ms and increases average fraud-detection rates by 20%, and by up to 300% in certain banks [15]. To comply with DORA and AMLD VI, scoring results are recorded in immutable logs, and the rules affecting the final risk score are made available to auditors via an Explainable AI layer, enabling regulators to reproduce the decision trajectory and verify the absence of discrimination.

The third scenario involves creditworthiness assessment. Large language models process unstructured data from application forms, call transcripts, and suppliers' ESG reports to generate features for a hybrid scoring system that combines a traditional probability-of-default module with an LLM-based text-quality evaluation. A McKinsey survey shows that already 20% of the banks have adopted at least one generative-AI use case in credit risk, and another 60% are going to adopt it within a year, as shown in Figure 3 [16].

Under the AI Act and FERPA, such systems being considered high-risk makes banks document their data sources, perform deviation outlier tests, and store each version of the model in a repository that can link up exactly which model configuration and feature set was used for a given credit decision. Thus, personalized recommendations, anti-fraud measures, and scoring come together into one unified technological stack sharing tracing mechanisms across the board; this drastically reduces the regulatory compliance cost as an added benefit of enabling go-lives for new AI functions.

The integration of AI into educational products logically continues the previous analysis: clients still want measurable algorithmic impact, regulators still require transparency, and developers still need to balance fast experimentation with keeping training data confidential. Against this backdrop, demand for flexible and explainable models is growing rapidly in both corporate training and the public sector, since these models enable the combination of personalized learning with requirements for personal data protection and minimization of discriminatory risk [17].

The scale of AI usage in schools is already comparable with the corporate segment: a survey by the Center for Democracy & Technology showed that in the 2023/24 academic year, 70% of high-school students and 67% of teachers regularly used generative tools, compared to 58% and 51%, respectively, the year before. Among the most in-demand solutions are adaptive learning platforms and chat-tutors that generate a recommendation stream of exercises based on each student's current level and errors [18].

The high degree of personalization reinforces the need for early detection of academic risk. As shown by Latin American practice, government analytics dashboards that combine attendance, grades, and LMS behavior signals provide administrators up to four weeks to take preventative action, which lessens the burden on academic counselors [17]. Finally, the expansion of AI's creative capabilities has intensified concerns about academic integrity. According to Turnitin, out of more than 200 million submitted works in 2024, signs of machine-generated text were detected in approximately 11% of cases, with only 3% of essays being almost entirely generated by AI [19]. Simultaneously, 39% of educators already use detection tools regularly; however, experts note the risk of false positives and emphasize the need for explainability of these systems' conclusions [18]. The advent of synthetic pictures and video further complicates matters: over the last year, deep-fake content observed in America grew by 303%, as per Ole Miss, thereby forcing organizations to update their cybersecurity strategies once again and now include fast media-content authentication steps [20]. This is, therefore, what makes auditing training datasets plus keeping request logs important because without these, there is no way to prove to regulators that data processing and model outputs do not violate fairness and proportionality principles.

Effective regulatory compliance begins with risk mapping and model classification; developers match target functions to regulator categories, define permissible automation levels, and record all assumptions in a registry accessible to auditors. Such early structuring immediately distinguishes high-risk scenarios—such as credit scoring or automated knowledge assessment—from low-risk auxiliary functions, thereby setting the appropriate depth for subsequent audits.

Next comes data governance based on the principles of minimization, provenance tracking, and formal reporting: teams document each attribute, justify its necessity, implement pseudonymization processes, and remove unnecessary fields before the training phase. Metadata captures end-to-end lineage; this means being able to rebuild the entire data path from its source up to model features, thus supporting incident investigations. The third layer is one of transparency and explainability. Users have to understand the logic behind recommendations or rejections. At the same time, regulators need sufficient information for verification, thus creating an interpretation layer in architecture where, for every decision made, the distribution of feature contributions is stored and presented in a human-readable form. These reports are embedded within the client or student interface and simultaneously exported into the control system- thereby fulfilling requirements from both sectors.

Regular checks help keep the math fair. Stats for soft spots get tallied during setup and run; if numbers go above set limits, the model gets more learning time or new weight tweaks to steer in the right way. A stand-alone ethics board stays clear of the product crew, cutting out any conflict of interest, and their check makes it into an open write-up.

This is ML-Ops under compliance. All model versions go through and pass the two-stage approval comparator process of technical metrics against the previous release, and yes, by the compliance officer who checks all explanatory artifacts. In case any of a checklist of conditions fails, auto-rollback happens, thereby ensuring that no risk-critical change reaches production without formal approval.

Real-time streaming metrics shall be computed with routing alerts to the central response center upon anomaly detection, switching the system's traffic to the fallback model or even manual review if necessary. This kind of closed loop provides resilience and cyber-reliability, enabling scaling up AI service by financial and educational organizations without getting penalized or even losing users' trust.

Continuing the logical chain of risk management, data protection techniques form the practical foundation on which trust in algorithms is built. The first layer is pseudonymization: as soon as the data is ingested, actual identifiers are immediately replaced with persistent hashes, and the key is disentangled only on an isolated node so that analysts can work with statistical profiles rather than personal data. This reduces the attack surface yet maintains the possibility to contact a user if the model finds an anomaly or when it issues a recommendation.

Wherever by law, the passing of raw features across borders of jurisdictions is not allowed, federated learning models will be used. The algorithm gets trained on devices or local clusters directly; only encrypted gradients are sent to the central orchestrator, thereby minimizing the cost of cross-border data processing agreements. It is more robust, too, because the failure of one participant cannot stop the collective parameter-update process.

To prevent reconstruction of original data from intermediate statistics, mathematically calibrated noise is added to aggregated results by differential privacy principles. The key advantage of this method is the strict guarantee that publishing model outputs will not increase the probability of disclosing any individual's attributes, thereby allowing researchers to share summaries without compromising confidentiality. However, increased noise entails higher error, so organizations define an acceptable privacy budget and record it in their secure-publication policy.

Synthetic data, created from the real distribution but not directly related to real individuals, is also used. This permits augmenting the training sample for rare classes, running stress tests on the anti-fraud system, and also sharing data between departments, bypassing lengthy legal procedures. The key thing here is to ensure that synthetic data does not repeat the original records using their unique combination. For this purpose, a couple of metrics are used to check diversity as well as the distance between samples.

Technical measures remain effective only with institutional support practices. Financial organizations publish an open manifesto listing all operational models, describing their purpose, key metrics, and emergency-shutdown procedures; the document is signed by the chief risk officer and updated according to a defined schedule. This format makes processes predictable for regulators and transparent to clients, while also reducing the likelihood of redundant internal controls.

In the educational sector, the resilience of the ecosystem is guaranteed by AI ethics committees. Membership draws from educators and administrators as well as legal experts and students, permitting consideration of academic, legal, and social aspects. The purpose of algorithm deployment, transparency requirements, thresholds for intervention, and policies on data retention are reviewed by the committees, and thereafter, binding recommendations to be implemented in the learning environment are issued.

The final link comprises external auditors and the practice of open reporting. Independent experts verify model compliance with declared policies, examine control samples for bias, and confirm the accuracy of documentation. Audit results are published publicly, providing societal oversight and fostering continuous improvement.

Collectively, these processes form a closed-loop cycle of trust in which technical solutions, organizational structures, and public reporting mutually reinforce one another, allowing the scaling of artificial intelligence without compromising user rights or business resilience.

IV. Conclusion

In conclusion, implementing artificial intelligence in financial and educational products requires a comprehensive approach that combines technical, organizational, and regulatory mechanisms. Analysis of European Union and U.S. legislative initiatives revealed requirements that are similar in substance yet different in form regarding risk classification, algorithmic explainability, and data protection. In both domains, three vectors proved central: precise identification of high-risk scenarios, transparency of decisions, and strict minimization of processed data. This enables consolidated event logging, bias auditing, and real-time model drift response at the technical layer, being compliant with the AI Act, PSD2, DORA, FERPA, and COPPA, as well as going far above it. The technological architecture presented herein is the materialization of this framework in practice: starting from a feature-store with mandatory logging up to an Explainable AI layer towards streaming quality monitoring, including automatic fallback to reserve models. Pseudonymization, federated learning, differential privacy, and synthetic data generation ensure functional compliance with jurisdictional requirements, enabled by practical means, so that there is a minimized possibility of leaks or misuse. At the same time, it ensures anomaly detection and correction using ML-Ops metrics in different layers and emergency shutdown procedures enabled by practical means, so that there is a minimized possibility of leaks or misuse.

Organizational practices sit well with technical safeguards: open model manifestos inside banking institutions and AI ethics committees in educational establishments make interaction between developers, administrators, and external auditors transparent. It not only helps to facilitate regulatory compliance but also builds up a culture for responsible AI deployment wherein every step from design up to production release is based on documented protocols and is under the scrutiny of an independent review.

This article unveils a cycle of trust wherein new AI solutions are implemented based on risk mapping and ready data governance, mandatory explainability, and oversight in place. Implementing these practices enables the scaling of artificial intelligence capabilities without infringing upon user rights or business stability, thereby striking a balance between technological progress and adherence to the highest standards of security and ethics.

References

[1] A. Singla, A. Sukharevsky, L. Yee, M. Chui, and B. Hall, “The state of AI: How organizations are rewiring to capture value,” *McKinsey & Company*, Mar. 12, 2025.

https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai (accessed Jun. 31, 2025).

[2] C. Ji, “3 parts of the market where AI hype is translating into real returns,” *Business Insider*, Jul. 24, 2025. https://www.businessinsider.com/ai-hype-translating-real-returns-investors-financials-morgan-stanley-2025-7 (accessed Jun. 31, 2025).

[3] “2025 AI in Education,” Microsoft, 2025. Accessed: Jul. 01, 2025. [Online]. Available: https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/bade/documents/products-and-services/en-us/education/2025-Microsoft-AI-in-Education-Report.pdf

[4] A. M. Ash, “Three in 10 Teachers Use AI Weekly, Saving Six Weeks a Year,” *Gallup*, Jun. 25, 2025. https://news.gallup.com/poll/691967/three-teachers-weekly-saving-six-weeks-year.aspx (accessed Jul. 02, 2025).

[5] European Commission, “AI Act,” *European Commission*, Feb. 18, 2025. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (accessed Jul. 03, 2025).

[6] EBA, “Opinion on new types of payment fraud and possible mitigants,” EBA, 2024. Accessed: Jul. 04, 2025. [Online]. Available: https://www.eba.europa.eu/sites/default/files/2024-04/363649ff-27b4-4210-95a6-0a87c9e21272/Opinion%20on%20new%20types%20of%20payment%20fraud%20and%20possible%20mitigations.pdf

[7] “EU 2022/2554,” *DORA*. https://www.digital-operational-resilience-act.com/ (accessed Jul. 04, 2025).

[8] “The new European AML package,” Fresfields. Accessed: Jul. 05, 2025. [Online]. Available: https://www.freshfields.com/globalassets/noindex/documents/the-european-aml-package-a-navigator.pdf

[9] “GDPR Enforcement Tracker - list of GDPR fines,” *Enforcement tracker*. https://www.enforcementtracker.com/?insights= (accessed Jul. 06, 2025).

[10] CDC, “Family Educational Rights and Privacy Act (FERPA),” *Public Health Law*, May 13, 2024. https://www.cdc.gov/phlp/php/resources/family-educational-rights-and-privacy-act-ferpa.html (accessed Jul. 07, 2025).

[11] “FTC proposes strengthening children’s online privacy rules to address tracking, push notifications,” *AP News*, Dec. 20, 2023. https://apnews.com/article/ftc-children-social-media-games-coppa-352ba63293832ee930f0c137aac735de (accessed Jul. 08, 2025).

[12] “America’s AI Action Plan: What’s In, What’s Out, What’s Next,” *HK Law*, 2025. https://www.hklaw.com/en/insights/publications/2025/07/americas-ai-action-plan-whats-in-whats-out-whats-next (accessed Jul. 08, 2025).

[13] D. Ivanov and D. Iaskova, “Generative AI in Banking: Practical Use Cases and Future Potential,” *Trinetix*, Jun. 06, 2024. https://www.trinetix.com/insights/generative-ai-in-banking (accessed Jul. 09, 2025).

[14] Visa, “Biannual Threats Report,” Visa, 2025. Accessed: Jul. 09, 2025. [Online]. Available: https://corporate.visa.com/content/dam/VCOM/corporate/solutions/documents/visa-perc-biannual-report-spring-2025.pdf

[15] “Mastercard supercharges consumer protection with gen ai,” *Mastercard*, 2024. https://www.mastercard.com/us/en/news-and-trends/press/2024/february/mastercard-supercharges-consumer-protection-with-gen-ai.html (accessed Jul. 10, 2025).

[16] McKinsey & Company, “Embracing generative AI in credit risk,” *McKinsey & Company*, Jul. 01, 2024. https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/embracing-generative-ai-in-credit-risk (accessed Jul. 11, 2025).

[17] “What You Need to Know AI Revolution in Education,” World Bank, 2025. Accessed: Jul. 11, 2025. [Online]. Available: https://documents1.worldbank.org/curated/en/099734306182493324/pdf/IDU152823b13109c514ebd19c241a289470b6902.pdf

[18] A. Merod, “Student, teacher AI use continued to climb in 2023-24 school year,” *K-12 Dive*, Jan. 15, 2025. https://www.k12dive.com/news/student-teacher-ai-use-schools-cdt/737335/ (accessed Jul. 12, 2025).

[19] A. Prothero, “New Data Reveal How Many Students Are Using AI to Cheat,” *Education Week*, Apr. 25, 2024. Accessed: Jul. 14, 2025. [Online]. Available: https://www.edweek.org/technology/new-data-reveal-how-many-students-are-using-ai-to-cheat/2024/04

[20] “Researchers Study How Deepfake Detection Tools Influence Journalists,” *Olemiss*, 2023. https://olemiss.edu/news/2024/07/dungeons-and-deepfakes/index.html (accessed Jul. 16, 2025).