International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Federal vs. Platform Responsibility in Preventing Digital Fraud: The Case of Online Dating and FinTech Apps

Inna Simonova

Point Break Capital, Associate (Business Consulting), Riverside, USA
Email: *inna.simonova[at]hotmail.com*

Abstract: ***In 2024, the convergence of social interaction platforms and financial technologies became a trigger for a large-scale crisis of digital trust, manifested in the de facto industrialization of romantic fraudulent practices and the increased complexity of abuse models associated with chargebacks. The conducted analysis records a shift in emphases in legal regulation and, as a consequence, in the construction of legal liability: the traditionally strong protective framework of the immunity of Section 230 of the Communications Decency Act (USA) is compared with a regulatorily opposite model in its logic, namely a proactive duty of care institutionalized in the United Kingdom Online Safety Act 2023 (OSA) and the EU Digital Services Act (DSA). Based on an interpretation of fraud statistics for 2024, reflecting a 25% increase in consumer losses to $12.5 billion, and incorporating the case of Social Discovery Group (SDG), the conclusion is substantiated that the automation of dispute resolution procedures and the application of machine learning tools have ceased to be merely a means of operational optimization and have transformed into the only normatively justified protection mechanism under conditions of pronounced regulatory turbulence. The obtained results indicate that the deployment of a Compliance-by-Design architecture provides platforms with a 33% reduction in operational costs for chargeback processing and forms resilience to adversarial impacts on AI models, including targeted data poisoning.***

Keywords: Digital fraud, Platform liability, FinTech, Online dating, Friendly Fraud, Pig Butchering, Section 230 CDA, Dispute automation, Machine learning, Adversarial attacks.

1. Introduction

The contemporary digital economy has encountered a phenomenon that the expert community describes as a fraud pandemic: as the boundaries between social interaction and financial transactions are erased, an environment conducive to hybrid threats is being formed. Statistics of the U.S. Federal Trade Commission (FTC) indicate that in 2024 the aggregate consumer losses from fraudulent practices reached a historical peak of $12.5 billion, exceeding the previous year's level by 25% with a virtually unchanged number of complaints [1]. This divergence, a stable volume of reports against the backdrop of a sharp increase in monetary damage, points not to an expansion of reach but to a qualitative strengthening of attacking vectors and an increase in the monetizability of each incident.

Assessments by Cybersecurity Ventures complement the picture of systemic risk, projecting that by 2025 global damages from cybercrime will rise to $10.5 trillion per year, which corresponds to one of the largest transfers of economic benefits in the entire history of observation [3]. Within this macrodynamic, the most vulnerable contours manifest in segments located at the intersection of dating and financial services. The median damage in investment romance scams exceeded $9 000, and the involvement of cryptocurrencies as an instrument of theft reached $1.4 billion, which reflects the institutionalization of highly liquid channels for the withdrawal of funds.

Alongside criminal fraud, pressure from friendly fraud is also increasing, in which legitimate users initiate disputes of their own transactions, transforming the payment infrastructure into a field of transactional conflicts. In 2024, 72% of merchants in e-commerce reported an increase in such incidents, which scales operational costs and increases the load on support and risk-control contours [4]. Against the backdrop of an increase in the cost of countering fraud for financial institutions in North America to $5.75 for each dollar lost, 5 manual moderation practices and reactive response models demonstrate structural infeasibility, as they do not correspond to the speed, adaptability, and economics of contemporary fraud ecosystems.

The objective of the work is manifested in identifying how competing legal models of liability allocation (immunity under Section 230 in the United States vs. duty of care in the United Kingdom and the EU) transform requirements for countering digital fraud at the junction of online dating and FinTech, and in substantiating the necessity of transitioning to a Compliance-by-Design architecture with dispute automation and ML detection.

Scientific novelty consists in proposing an interdisciplinary analysis that links the evolution of regulatory liability regimes with measurable operational effects (chargebacks, win rate, false positives, costs) and new classes of threats (LLM-enhanced fraud, deepfakes, adversarial AI/data poisoning), showing that automation and AI become not an optimization option but a normatively justified mechanism for compliance with reasonable procedures.

Practical **significance** is reduced to the assumption that the results provide an applied basis for platforms and payment participants to implement a reproducible compliance contour (ADR, hybrid AI, feedback based on dispute outcomes, explainability, and resilience to adversarial impacts), enabling a simultaneous reduction of losses and operational costs (including those for chargeback processing) and the

demonstration of verifiable compliance with OSA/DSA requirements and with the discussed models of joint liability.

2. Materials and Methods

The study is based on a mixed methodological design combining the instruments of legal hermeneutics, statistical processing of large data arrays, and technical modeling of applied cases, which ensures the comparability of legal constructs with observable operational effects and technological constraints.

The empirical framework is formed on the material of the FTC Consumer Sentinel Network Data Book 2024 reports [6], the annual analytical studies LexisNexis True Cost of Fraud [5], as well as the technical documentation of the payment systems Visa and Mastercard relating to the update of dispute reason codes in 2024 [8]. The regulatory and legal block is interpreted on the basis of the texts of acts of the United States (Communications Decency Act, SAFE TECH Act), the United Kingdom (Online Safety Act 2023), and the European Union (Digital Services Act, Payment Services Regulation), which makes it possible to trace differences in the logic of liability allocation, due diligence standards, and requirements for preventive compliance.

The key applied module of the study is built around the operational model of Social Discovery Group (SDG) as a representative example of a platform ecosystem functioning at the intersection of social interaction and payment infrastructure. The analysis of this case is conducted through an assessment of the implementation of automated dispute resolution systems (ADR) and the application of generative AI for anomaly detection, with a focus on measurable technical indicators. As critically significant metrics, the reduction of the share of false positive activations and the chargeback dispute win rate are considered, which on average across the industry is only 45% [10], yet is amenable to substantial improvement when transitioning to algorithmic processing of evidence and standardization of the evidentiary base within machine-oriented decision-making contours.

3. Results and Discussion

The global digital environment demonstrates increasing fragmentation of legal regimes, as a result of which cross-border platforms are compelled to align their risk management models with competing normative logics: the American doctrine of limited liability and the European-British imperative of preventive due diligence, which codifies the obligation of active harm prevention.

In the United States, the foundation of regulation for a long time was Section 230 of the Communications Decency Act (Section 230 CDA), providing platforms with immunity from liability for user content. However, by 2024 a trend toward rethinking the boundaries of this immunity emerged. Despite the Supreme Court's refusal to consider the case Doe v. Grindr on the merits, the separate opinions of Justices Thomas and Gorsuch are interpreted as an indicator of readiness to narrow the protective perimeter of Section 230, primarily in situations where the platform's recommendation and ranking mechanisms functionally facilitate the commission of unlawful acts [11]. In parallel, a line of liability is strengthening that is based not on attributing third-party content, but on qualifying the platform's own practices as misleading: the settlement of the Federal Trade Commission case against Match Group with a payment of $14 million formed an illustrative example in which the subject of claims was a complicated subscription cancellation procedure and the exploitation of fake profiles in advertising scenarios (love interest ads) [13]. Against this background, the SAFE TECH Act bill, associated with Senator Warner's initiative, is aimed at excluding the application of Section 230 immunity for cases related to paid content and fraudulent advertising, which conceptually transforms the risk profile of dating platforms and other services monetized through subscriptions and targeting [15]. An additional vector of institutionalization of oversight is formed through the Algorithmic Accountability Act, providing for the obligation of companies to conduct impact assessments of automated decision-making systems on critical spheres of consumer life, including financial services [17, 23].

The United Kingdom, by contrast, consistently enshrines a normative model of prevention as a baseline standard. The Online Safety Act 2023 (OSA) qualifies fraud as a priority offense, thereby shifting it from the domain of ex post response to a regime of mandated proactive risk management. For Category 1 platforms, pertaining to the largest services, an obligation is established to implement technical and organizational measures aimed at preventing the emergence and dissemination of fraudulent content, including the segment of paid advertising [18]. At the same time, the personal-organizational measurability of compliance is strengthened: the introduction of the offense Failure to Prevent Fraud for large organizations implies the possibility of liability for the actions of employees or agents in the absence of demonstrable reasonable procedures of prevention [19]. Thus, the due diligence standard acquires the character of a procedurally verifiable system of measures, rather than a declarative obligation.

The European Union develops a model of distributed, functionally interconnected responsibility, combining requirements for transparency of the platform economy with the strengthening of the payment perimeter. The Digital Services Act (DSA) sets out strict disclosure obligations and introduces Know Your Business Customer (KYBC) requirements for marketplaces, forming a regulatory infrastructure for cutting off commercial abuses at the stage of access to an audience [20]. The most transformative consequences, however, are associated with discussions around reforming the Payment Services Regulation (PSR/PSD3): the European Parliament is considering a mechanism of joint and several liability, entailing a financial obligation for online platforms and electronic communications service providers (ECSP) to participate in compensating losses to victims of APP fraud (Authorized Push Payment) in the event of an inability to prevent the dissemination of fraudulent content or to block it in a timely manner [21, 22]. Such a construct forms a direct economic incentive to invest in detection systems and reduces the attractiveness of minimal-compliance strategies, since the cost of inaction acquires a measurable, compensatory character.

Below, in Table 1, for greater clarity, the results of the comparative analysis of regulatory liability regimes will be presented.

Table 1: Comparative analysis of regulatory liability regimes (compiled by the author based on [15, 21, 22]).

Characteristic	USA (Section 230 / SAFE TECH)	United Kingdom (OSA 2023)	EU (DSA / PSR Proposals)
Basic principle	Intermediary immunity (with exceptions).	Duty of Care.	Intermediary liability and transparency.
Liability for fraud	Limited (if the platform is not the creator).	Strict (Failure to Prevent).	Joint and several (proposed allocation with banks).
Verification requirements	Voluntary / patchwork regulation across states.	Mandatory (child protection).	Mandatory KYBC for marketplaces.
Enforcement mechanism	Civil lawsuits, FTC (consumer deception).	Ofcom (fines up to 10% of turnover).	European Commission / Digital Services Coordinators.
Status 2024	Judicial precedents (Grindr, Match), draft bills.	Implementation phase.	In force; PSR under discussion.

The regulatory requirement to implement reasonable procedures enters into a direct contradiction with the accelerating technological evolution of fraudulent practices. By 2024, the image of the ideal fraudster had lost its former relevance: instead of demonstratively flawless narratives, the tactic of strategic imperfection became established, within which trust is constructed not through idealization, but through controlled plausibility and the reduction of signals associated with automated spam detection.

Pig Butchering schemes evolved into high-technology operations with features of industrial production. To scale attacks, large language models (LLM) are used, enabling the conduct of dialogues that are linguistically difficult to distinguish from native speakers' speech, which increases the throughput of offenders without a proportional increase in personnel costs [16, 24]. The adaptation vector is also manifested at the level of social engineering: a study by the University of Missouri recorded a shift in the paradigm of profile construction, in which instead of glossy images with an abnormally high social capital, everyday flaws are deliberately embedded into narratives (divorce, ordinary employment, moderate life difficulties). Such normalization of the profile reduces the probability of triggering filters oriented toward identifying excessively successful personas and simultaneously strengthens the trust effect due to psychologically recognizable vulnerability [14, 25].

In parallel, a sharp intensification of attacks based on synthetic media is observed. The use of deepfakes in the financial sector increased by 700% [26], and an illustrative incident in 2024, in which an employee of a Hong Kong company executed a transfer of $25 million after a video call with a deepfake of the chief financial officer, demonstrates the systemic vulnerability of traditional channels of verification and authorization. At the level of forecasts, it is indicated that in 2025 deepfakes may account for 5% of all fraudulent attacks, which is equivalent to an increase of 1300% compared with 2023 [27]. In such a configuration, static identity verification (IDV) loses sufficiency as an independent barrier and requires a shift toward dynamic methods, including behavioral biometrics and liveness analysis capable of detecting artificially synthesized representations [9, 12].

An underestimated but critically significant category of threats is adversarial attacks, in particular model poisoning. The mechanism is based on the introduction of noisy or specially labeled data fragments into training samples, which systematically disrupts the classifier's ability to generalize and shifts decision boundaries. Empirical findings indicate that the substitution of approximately 0,001% of tokens in a training set can initiate persistent model errors [28]. For the FinTech context, this translates into the risk of an increase in missed fraudulent transactions being treated as legitimate or, conversely, into service degradation under the load of false positive activations, which turns AI contours into an object of attack rather than exclusively a tool of protection.

Under conditions of simultaneous regulatory pressure and increasing threat complexity, process automation acquires the character not of optimization, but of a structurally necessary response. The operational model of Social Discovery Group (SDG) illustrates the transition from reactive protection to a predictive Compliance-by-Design architecture, in which compliance is embedded at the level of process and integration design rather than added ex post. Traditional chargeback processing is characterized by high labor intensity: manual preparation of an evidentiary base can take up to 45 minutes per case, whereas complex fraud episodes require up to three hours of an analyst's work [30]. Given an annual increase in chargeback volume of 15%, manual procedures lose economic viability and predictability of quality [31]. In response, SDG implements an automated dispute resolution (ADR) contour integrated with payment gateway APIs and a CRM system, where primary analytics is built around automatic parsing of the chargeback reason code and the corresponding logic for generating evidence. A substantial complicating factor is changes in payment system classifiers: in April 2024 Visa updated the Reason Codes structure, merging code 12.1 (Late Presentment) with 11.3 (No Authorization), which strengthens requirements for the precision of response routing rules and for the correct configuration of evidentiary presentation scenarios [32].

Figure 1 presents the conceptual ADR workflow.

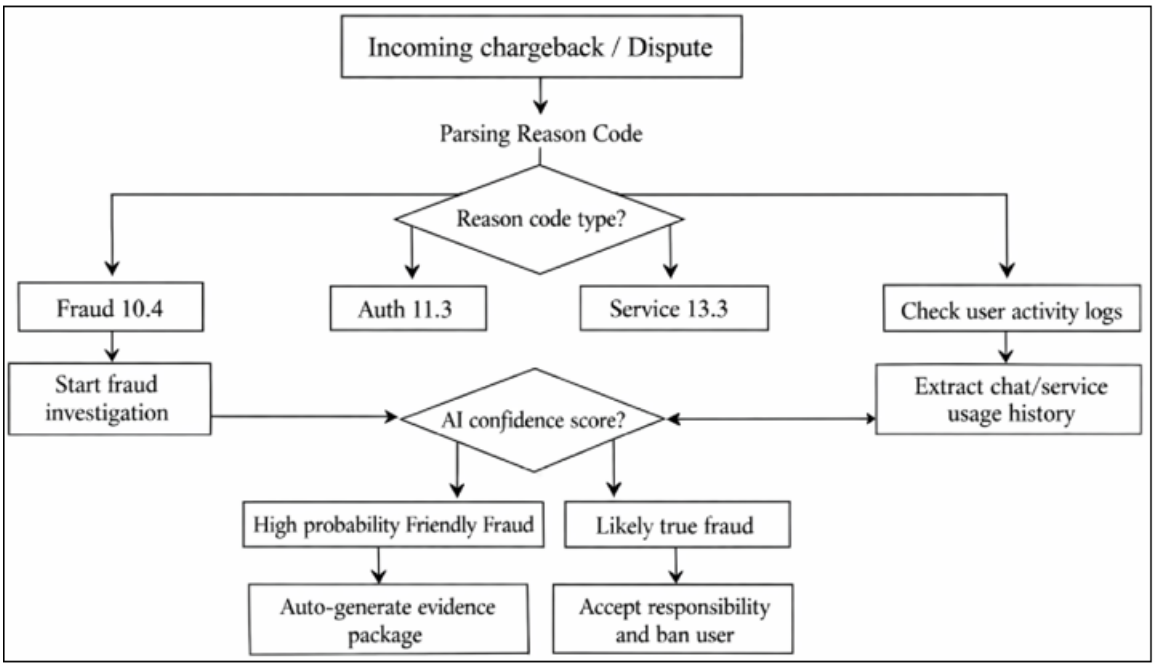

Figure 1: Conceptual workflow of ADR (compiled by the author based on [32, 33]).

Automation functions not only as a tool for reducing labor inputs, but also as a mechanism for ensuring the completeness of processing of provable episodes, making it possible to bring 100% of valid cases to the dispute stage without losses caused by the human factor and operational bottlenecks. Data from Verifi and Chargebacks 911 demonstrate that the use of automated responses in combination with the Compelling Evidence 3.0 rules increases the effectiveness of dispute resolution: merchants are able to win up to 45% of disputes, and in the digital goods segment, in the presence of reliable confirmations of actual consumption of the service, the indicator can reach 72% [10].

To comply with the United Kingdom Failure to Prevent construct, SDG employs machine learning models that functionally surpass static rule sets. Instead of primitive blocking by IP addresses, risk assessment is built on the analysis of behavioral trajectories and sequences of actions, which makes it possible to identify behavioral patterns that are resilient to changes in offenders' infrastructure. A characteristic example is Pig Butchering fraud, which demonstrates a specific speed signature: an accelerated transition from initial contact to imposing migration of communication to encrypted channels (WhatsApp/Telegram), as a rule accompanied by repeated trigger markers in the lexicon and interaction scenarios [2, 24].

Resilience to adversarial impacts, including data poisoning, is ensured through ensemble approaches and anomaly detection methods oriented toward identifying statistically atypical distributions in input streams characteristic of noise injections and targeted distortions of training samples [7, 29].

A high share of false positive activations remains one of the system-forming constraints of automation: on average across the industry, the indicator is estimated in the range of 2–10%, whereas in legacy antifraud contours it can exceed 10% [34]. Such a classification error translates into direct commercial losses due to blocking legitimate users, degradation of user experience, and foregone revenue. To neutralize this effect, SDG uses a Hybrid AI architecture in which supervised learning on historical labeled arrays is combined with unsupervised learning oriented toward detecting previously unknown anomalies and atypical behavioral clusters [35]. Such a combination increases the system's resilience to data drift and the emergence of new offender tactics, while simultaneously reducing dependence on rigidly specified rules.

The constructive core of this architecture is a feedback loop: dispute outcomes (won/lost) are returned to the model as a quality signal and a basis for revising classification boundaries. In cases where a chargeback is successfully disputed and qualified as Friendly Fraud, a training impulse is formed that makes it possible to separate criminal fraud from behavior associated with consumer extremism. As a result, the behavioral profile ceases to be interpreted as an indicator of criminal fraud and is transferred into the category of transactional risk requiring other response measures, for example, limiting refund procedures or changing payment parameters, but without automatic account blocking as a measure disproportionate to the nature of the detected violation.

An analysis of operational data for 2024 confirms that the transition to automated protection systems is not only a legal imperative, but also an economically justified strategy.

The financial assessment of aggregate damage associated with fraud is increasingly expressed through the fraud multiplier indicator, reflecting the magnitude of a company's losses per dollar of direct fraud. In 2024, the value of this indicator reached $5.75 [5]. A substantial share of the multiplier is formed not by the fact of theft itself, but by derivative costs, namely administrative costs for investigation, preparation of evidence, interaction with payment systems, and dispute support. Automation of representation in disputes reduces precisely this cost component, transferring case processing from a labor-

intensive mode, in which manual preparation can take about 45 minutes, into a high-speed contour performing key operations in seconds. An additional effect is achieved through technological measures at the payment level: the use of tokenization, according to industry data, correlates with a 15% reduction in the number of chargebacks on the grounds of unauthorized transactions in the CNP (card-not-present) environment [4].

At the same time, the scaling of AI approaches heightens ethical and legal risk of algorithmic bias, in which statistical regularities in data can lead to systematically unfavorable decisions with respect to particular demographic groups, which contradicts principles of fairness widely discussed in the academic literature [36, 37]. In the financial-technology context, such distortions can manifest in unjustified denials of service, asymmetric restriction of functionality, or disproportionate increases in risk scores. In response, SDG relies on explainable AI (Explainable AI / White Box) [35], which simultaneously reduces the probability of uncontrolled discrimination and strengthens normative compatibility with DSA requirements for the transparency of algorithmic mechanisms. The formation of human-readable reports accompanying disputes and system decisions effectively transforms the results of machine analysis into an evidentiary trajectory verifiable by auditors, in which the grounds for conclusions are reconstructed and reproducibility of oversight is ensured.

Table 2 reflects statistics of fraud volumes and financial impact.

Table 2: Statistics of fraud volumes and financial impact (compiled by the author based on [1]).

Metric	Value 2023	Value 2024	Change (YoY)	Implications for platforms
Total consumer losses (FTC)	$10,0 bn	$12,5 bn	25%	Increase in regulatory scrutiny and fines.
Median losses	~$2,000	~$2,000+ (Pig Butchering >$9k)	Stable/Growth	High value of each prevented incident.
Growth of Friendly Fraud	--	72% of merchants report growth	High growth	Need to automate evidence collection.
Dispute administration cost	--	$288 - $371 per dispute	Growth	Manual processing becomes unprofitable.
Chargeback win rate (Average)	45%	45%	Plateau	Static tools have exhausted their potential.
Use of deepfakes	--	Growth by 700%	Explosive growth	Crisis of trust in standard IDV.

An in-depth analysis makes it possible to identify a second-order effect manifested as a displacement of liability. As the protective contour of platforms is strengthened through AI tools and automated compliance, fraudulent activity is redistributed toward less protected participants of the ecosystem, namely services that still rely predominantly on the immunity doctrine of Section 230 and do not develop preventive mechanisms for detecting and suppressing abuse. As a result, a structural stratification of the market is formed, in which efforts to reduce risk in one part of the chain increase the relative vulnerability of another part, and offenders rationally migrate to where the costs of attack are lower and the probability of detection and sanctioning is minimal.

This shift creates a compliance moat, a competitive barrier based on the ability to maintain a normatively acceptable level of control and procedural demonstrability. Technologically mature companies prove more resilient under conditions of regulatory cleansing, including scenarios of increased oversight by the United Kingdom OFCOM, because they are able to demonstrate fulfillment of requirements for reasonable procedures, transparency, and risk management. By contrast, lagging actors face growing existential threats: legal claims intensify due to the absence of preventive measures, and operational viability is undermined by a possible loss of access to payment infrastructure if thresholds for fraud monitoring applied by Visa and Mastercard are exceeded, which moves the compliance problem from the plane of reputational costs into the category of an immediate risk of business cessation.

4. Conclusion

The era of permissive innovation in digital social spaces demonstrates signs of completion. The growth of fraud-related damage to $12.5 billion in 2024 served as a catalyst for regulatory tightening in the United Kingdom and the European Union, whereas in the United States a similar trajectory is beginning to manifest through judicial reinterpretations and the activation of legislative initiatives. Under these conditions, the former dichotomy of federal liability versus platform liability is consistently shifting in favor of the imperative of Platform Liability as the dominant normative logic.

The Social Discovery Group case shows that this liability is reduced not to a formal legal burden, but is transformed into an operational condition of viability. Automation of the dispute life cycle and the deployment of AI contours for detecting strategic imperfections of contemporary fraud scenarios form a triple-win effect: legal compatibility with the requirement of reasonable procedures under the United Kingdom OSA is achieved while simultaneously reducing regulatory risks within the DSA ecosystem; financial stabilization is ensured through the recovery of revenue lost due to Friendly Fraud and a 33% reduction in operational costs; the level of user safety is increased by disrupting Pig Butchering fraud chains before the withdrawal of funds from the controlled contour.

The prospective configuration of the industry is associated with a co-regulation model in which public standards codify a duty of care, and AI systems deployed on the platform side ensure its practical execution at scales inaccessible to manual procedures. Empirical observations indicate that manual compliance ceases to be a realistic alternative: the pace of innovation in the fraud environment requires an automated, algorithmically governed response capable of real-time adaptation and resilient to adversarial impacts, including attempts to manipulate data and degrade models.

References

[1] Federal Trade Commission. (2025, March 10). Top scams of 2024. Consumer Advice. Retrieved from: https://consumer.ftc.gov/consumer-alerts/2025/03/top-scams-2024 (date accessed: September 3, 2025).
[2] Federal Trade Commission. (2025, March 10). New FTC data show a big jump in reported losses to fraud to $12.5 billion in 2024. Retrieved from: https://www.ftc.gov/news-events/news/press-releases/2025/03/new-ftc-data-show-big-jump-reported-losses-fraud-125-billion-2024 (date accessed: September 5, 2025).
[3] Cybersecurity Ventures. (n.d.). Cybercrime to cost the world $10.5 trillion annually by 2025. Retrieved from: https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/ (date accessed: September 7, 2025).
[4] Visa Acceptance Solutions, Verifi, & Merchant Risk Council. (2025). 2025 Global eCommerce Payments & Fraud Report. Retrieved from: https://www.visaacceptance.com/content/dam/documents/campaign/fraud-report/global-fraud-report-2025.pdf (date accessed: September 9, 2025).
[5] LexisNexis Risk Solutions. (2025, September 10). Every dollar lost to fraud costs North America's financial institutions $5, according to LexisNexis Risk Solutions. Retrieved from: https://risk.lexisnexis.com/about-us/press-room/press-release/20250910-fraud-multiplier (date accessed: September 11, 2025).
[6] Federal Trade Commission. (2025, March). Consumer Sentinel Network Data Book 2024. Retrieved from: https://www.ftc.gov/system/files/ftc_gov/pdf/csn-annual-data-book-2024.pdf (date accessed: September 13, 2025).
[7] LexisNexis Risk Solutions. (2025, April 2). Fraud costs surge as North America's ecommerce and retail businesses face mounting financial and operational challenges. Retrieved from: https://risk.lexisnexis.com/about-us/press-room/press-release/20250402-tcof-ecommerce-and-retail (date accessed: September 15, 2025).
[8] Visa. (2024, June). Dispute Management Guidelines for Visa Merchants. Retrieved from: https://usa.visa.com/dam/VCOM/global/support-legal/documents/merchants-dispute-management-guidelines.pdf (date accessed: September 17, 2025).
[9] Mastercard. (2025, May 13). Chargeback Guide: Merchant Edition. Retrieved from: https://www.mastercard.us/content/dam/public/mastercardcom/na/global-site/documents/chargeback-guide.pdf (date accessed: September 19, 2025).
[10] Mastercard (Ethoca). (2025, March). The chargeback window of opportunity: A global view of the 2025 chargeback trends and how to turn them into opportunities. Retrieved from: https://www.mastercard.com/content/dam/mccom/shared/news-and-trends/insights/2025/2025-global-chargebacks-outlook/pdf/2025-state-of-chargebacks-report.pdf (date accessed: September 21, 2025).
[11] CBS News. (2024, July 2). Supreme Court rebuffs chance to evaluate scope of Section 230 legal shield in dispute involving Grindr. Retrieved from: https://www.cbsnews.com/news/supreme-court-section-230-grindr-case/ (date accessed: September 23, 2025).
[12] Supreme Court of the United States. (2024, July 2). Doe v. Snap, Inc., No. 23-961 (Opinion). Retrieved from: https://www.supremecourt.gov/opinions/23pdf/23-961_1924.pdf (date accessed: September 25, 2025).
[13] Federal Trade Commission. (2025, July 15). Proposed Order for Permanent Injunction, Monetary Judgment, and Other Relief (FTC v. Match Group, Inc.). Retrieved from: https://www.ftc.gov/system/files/ftc_gov/pdf/Match-Order.pdf(date accessed: September 27, 2025).
[14] Federal Trade Commission. (2025, August 20). Match Group agrees to pay $14 million, permanently stop deceptive advertising, cancellation, and billing practices to resolve FTC charges. Retrieved from: https://www.ftc.gov/news-events/news/press-releases/2025/08/match-group-agrees-pay-14-million-permanently-stop-deceptive-advertising-cancellation-billing (date accessed: September 29, 2025).
[15] Lawyers' Committee for Civil Rights Under Law. (2021). The SAFE TECH Act (Safeguarding Against Fraud, Exploitation, Threats, Extremism and Consumer Harms Act). Retrieved from: https://lawyerscommittee.org/wp-content/uploads/2021/02/SAFE-TECH-Act.pdf (date accessed: October 1, 2025).
[16] Congressional Research Service. (2023, May 19). Section 230: An overview (R46751). Retrieved from: https://crsreports.congress.gov/product/pdf/R/R46751 (date accessed: October 3, 2025).
[17] U.S. Congress. (2023). Text—H.R. 5628 (118th Congress): Algorithmic Accountability Act of 2023. Retrieved from: https://www.congress.gov/bill/118th-congress/house-bill/5628/text (date accessed: October 5, 2025).
[18] United Kingdom. (2023). Online Safety Act 2023. Legislation.gov.uk. Retrieved from: https://www.legislation.gov.uk/ukpga/2023/50 (date accessed: October 7, 2025).
[19] Bristows LLP. (2025, April). OSA vs DSA: Comparing the UK and EU online safety regimes (PDF). Retrieved from: https://www.bristows.com/app/uploads/2025/04/Comparing-the-UK-and-EU-online-safety-regimes.pdf (date accessed: October 9, 2025).
[20] European Commission. (2020, December 15). Questions and answers on the Digital Services Act. Retrieved from: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348 (date accessed: October 11, 2025).
[21] Centre for European Reform. (n.d.). Is the EU taking the right approach to APP fraud? Retrieved from: https://www.cer.eu/in-the-press/eu-taking-right-approach-app-fraud (date accessed: October 13, 2025).
[22] European Centre for International Political Economy. (n.d.). Shared liability: The European Parliament's misstep in fighting financial fraud. Retrieved from: https://ecipe.org/publications/ep-misstep-in-fighting-financial-fraud/ (date accessed: October 15, 2025).

[23] UK Government. (n.d.). Online Safety Act: Explainer. GOV.UK. Retrieved from: https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer (date accessed: October 17, 2025).
[24] Association of Certified Fraud Examiners. (2024). Report to the Nations 2024: Global study on occupational fraud and abuse. Retrieved from: https://legacy.acfe.com/report-to-the-nations/2024/ (date accessed: October 19, 2025).
[25] University of Missouri. (2025, February 10). Study shows online dating scammers' tactics are evolving. Show Me Mizzou. Retrieved from: https://showme.missouri.edu/2025/study-shows-online-dating-scammers-tactics-are-evolving/ (date accessed: October 21, 2025).
[26] Deloitte. (2025). Generative AI is expected to magnify the risk of deepfakes and other fraud in banking. Retrieved from: https://www.deloitte.com/us/en/insights/industry/financial-services/deepfake-banking-fraud-risk-on-the-rise.html (date accessed: October 23, 2025).
[27] World Economic Forum & Accenture. (2025). Artificial Intelligence in Financial Services. Retrieved from: https://reports.weforum.org/docs/WEF_Artificial_Intelligence_in_Financial_Services_2025.pdf (date accessed: October 25, 2025).
[28] Vassilev, A., & Oprea, A. (2025). Adversarial machine learning: A taxonomy and terminology of attacks and mitigations(NIST AI 100-2e2025). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-2e2025. Retrieved from: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=959735 (date accessed: October 27, 2025).
[29] Kure, H. I., Sarkar, P., Ndanusa, A. B., & Nwajana, A. O. (2025). Detecting and preventing data poisoning attacks on AI models. arXiv. https://doi.org/10.48550/arXiv.2503.09302. Retrieved from: https://arxiv.org/abs/2503.09302 (date accessed: October 29, 2025).
[30] Visa. (n.d.). Visa Resolve Online (VROL). Retrieved from: https://usa.visa.com/solutions/post-purchase-solutions/visa-resolve-online.html (date accessed: October 31, 2025).
[31] Chargebacks911. (2024). Chargeback Field Report 2024 (PDF). Retrieved from: https://ad.chargebacks911.com/hubfs/Chargeback%20Field%20Report%202024.pdf?_hsenc=p2ANqtz-8uaoNVRga4UzbJpfytSMC5Z0H2jwUwe7q73iZitNmkyww3yKHYJseBNxrpm_D7DYWOKVXXDXhXt5Cs9OyEkCPya5U9MMX_K2nOgJYL_JFsghoHvQM&_hsmi=79116209&utm_content=79116209&utm_medium=email&utm_source=hs_automation (date accessed: November 2, 2025).
[32] Visa. (2025). Visa Core Rules and Visa Product and Service Rules (Public Version) (PDF). Retrieved from: https://usa.visa.com/content/dam/VCOM/download/about-visa/visa-rules-public.pdf (date accessed: November 4, 2025).
[33] Visa. (2023, October 14). Introduction of monitoring rule for dispute condition 10.4: Other fraud—card-absent environment remedy (PDF). Retrieved from: https://corporate.visa.com/content/dam/VCOM/regional/na/us/support-legal/documents/introduction-of-monitoring-rule-for-dispute-condition-10.4-other-fraud-card-absent-environment-remedy.pdf (date accessed: November 6, 2025).
[34] Verifi. (2024). 2024 Global Fraud & Payments Report (PDF). Retrieved from: https://www.verifi.com/_assets/VERIFI_2024_Global-Fraud_Payments_Report.pdf (date accessed: November 8, 2025).
[35] NICE Actimize. (n.d.). Fraud detection analytics: A vital investment. Retrieved from: https://www.niceactimize.com/fraud-management/fraud-analytics-optimization/ (date accessed: November 10, 2025).
[36] Al-Daoud, K. I., & Abu-AlSondos, I. A. (2025). Robust AI for financial fraud detection in the GCC: A hybrid framework for imbalance, drift, and adversarial threats. Journal of Theoretical and Applied Electronic Commerce Research, 20(2), 121. https://doi.org/10.3390/jtaer20020121.
[37] Kamalaruban, P., Pi, Y., Burrell, S., Drage, E., Skalski, P., Wong, J., & Sutton, D. (2024). Evaluating fairness in transaction fraud models: Fairness metrics, bias audits, and challenges. In Proceedings of the 5th ACM International Conference on AI in Finance (ICAIF '24) (pp. 555–563). https://doi.org/10.1145/3677052.3698666.