Journal of Global Economy, Business and Finance (JGEBF) ISSN: 2141-5595
DOI: 10.53469/jgebf.2025.07(06).13

Shifting Liability Principles for Generative Artificial Intelligence from a Law and Economics Perspective: From Negligence Liability to Strict Liability

Chunxiao Mao

School of Law, Anhui University of Finance & Economics, Bengbu 233030, Anhui, China

Abstract: ***While generative artificial intelligence's technological breakthroughs unleash tremendous value, its unique tort risk structure fundamentally challenges traditional negligence liability principles. This thesis systematically analyzes, from a law and economics perspective, the inherent defects and institutional failure origins of applying negligence liability principles to Generative AI infringements, and demonstrates the rationality and advantages of shifting toward strict liability principles. Through typological analysis of Generative AI infringement scenarios, four core characteristics are distilled: technological black box and causal concealment, systemic risk diffusion, web-like dissolution of responsible entities, and inevitable damages with high remediation costs. The thesis indicates that negligence liability principles face systemic failures including judicial difficulties in determining reasonable care standards, excessive expansion of behavioral standards, and inherent defects in alternative attribution schemes. In contrast, strict liability principles demonstrate significant advantages by simplifying attribution requirements, optimizing care level incentives, and internalizing activity level costs. Therefore, this thesis argues that under existing legal frameworks and technological conditions, strict liability principles represent a more feasible and necessary institutional choice.***

Keywords: Generative artificial intelligence tort, Law and economics analysis, Strict liability principles, Causal concealment, Systemic risk diffusion.

1. Introduction

Generative artificial intelligence, as a phenomenal technological breakthrough, is reshaping social production and lifestyle paradigms with unprecedented breadth and depth. While its powerful content creation and information processing capabilities release enormous economic and social value, they simultaneously engender novel and complex tort risks. Generative AI infringement scenarios now extensively encompass core public interest domains including personality rights, intellectual property, and derivative torts. Its distinctive characteristics that differentiate it from traditional torts—technological opacity and causal concealment, systemic risk propagation, the networked dissolution of responsible entities, and the irreversibility of damages coupled with high remediation costs—not only intensify the complexity of damage assessment [1] but directly constitute structural impediments to the application of negligence liability, undermining the core attribution principle of negligence determination in generative artificial intelligence tort liability. The "reasonable duty of care" standard upon which negligence liability relies faces a dual predicament of judicial determination difficulties and escalating institutional costs within the context of algorithmic inexplicability, rapid technological iteration, and multi-stakeholder accountability dynamics, rendering its attribution logic systematically dysfunctional. More critically and fundamentally, the institutional design of negligence liability fails to effectively constrain excessive behavioral expansion; instead, by externalizing residual accident costs to society after reasonable care is taken, it causes risk production to far exceed socially optimal levels, thereby institutionalizing negative externalities as a consequence.

In this context, this thesis aims to systematically analyze the inherent defects and institutional failure origins of applying the negligence liability principle to generative artificial intelligence torts from a law and economics perspective, while demonstrating the institutional rationality and advantages of shifting toward strict liability principles. The law and economics analytical framework is particularly suitable for deconstructing liability challenges posed by such complex technologies, with its core advantage being the ability to penetrate technological appearances and employ cost-benefit analysis methods to clearly reveal decision-making logic of actors regarding "level of care" and "level of activity" under different attribution principles, as well as the resulting social costs, thereby providing efficiency-based rational foundations for institutional design [2]. This thesis will delve into several key issues: how strict liability principles can simplify attribution events, effectively circumventing the challenges of determining reasonable care and multi-party attribution games caused by technological black boxes; how risk internalization mechanisms can simultaneously optimize both the level of care and activity level of actors, guiding them to spontaneously converge toward socially optimal risk activity scales; and how the predetermined responsible parties and supporting mechanism design can ensure victim remedies and enhance risk control efficiency while avoiding undue suppression of technological innovation. This research aims to reveal that, in the face of generative artificial intelligence's unique tortious risk structure, strict liability is not only a more efficient attribution pathway, but also a more feasible and necessary institutional choice within the existing legal framework and technological conditions.

2. Structural Characteristics of Generative Artificial Intelligence Tort

The intrinsic properties of generative artificial intelligence fundamentally reshape the form and attribution basis of tortious conduct. To systematically deconstruct its liability issues, this section first provides a structured classification of

typical infringement scenarios involving Generative AI based on the types of legal interests harmed. It then distills four core characteristics that distinguish it from traditional torts, thereby establishing a theoretical foundation for subsequent analysis of the institutional compatibility of attribution principles.

2.1 Typology of Characteristic Infringement Scenarios

Generative AI infringement scenarios exhibit three distinctive features: technological endogeneity, diffusive harm, and ambiguity of responsible entities [3]. The tortious outcomes stem from algorithmic autonomous evolution during data training and content generation, rather than direct human behavioral instructions. Based on the types of legal interests harmed, these primarily encompass three categories of scenarios: personality Rights Infringement, intellectual property rights violation, and indirect tort.

2.1.1 Personality Rights Infringement Scenarios

Generative AI's infringement on personality rights primarily manifests as new risks including privacy breaches, reputation damage, unauthorized use of portraits, and harm to the personality interests of the deceased. In the domain of privacy and personal information, Large Language Models rely on massive datasets for training, and developers' unauthorized collection or utilization of personal data without consent constitutes direct tort upon the right to informational self-determination. From a technical perspective, model inversion attacks can extract sensitive information from training data, with attack success rates positively correlated with model scale. A more insidious risk lies in Generative AI's capability to infer undisclosed personal sensitive information through associative reasoning, such as sexual orientation and health status, effectively circumventing the "informed consent" principle [4]. Defamation liability arises from algorithmic hallucination characteristics, where autonomously generated false information — such as fabricated academic misconduct or criminal scandals — possesses high verisimilitude, with propagation velocity and remediation costs far exceeding traditional defamatory content. In the domain of publicity rights, Generative AI has transcended technical limitations of deepfakes, synthesizing dynamic imagery from single photographs, resulting in commercial misappropriation that causes identity confusion and image distortion. Regarding posthumous personality interests, algorithms aggregate deceased individuals' data to generate digital simulacra with personality continuity, producing discriminatory statements or fraudulent interactions, while current legal frameworks exhibit significant deficiencies in protecting the dignitary aspects of digital estates [5].

2.1.2 Intellectual Property Infringement Scenarios

Generative AI intellectual property infringement spans both the training data input and content generation output phases, with the core controversy centered on the ambiguity of "fair use" boundaries in copyright law. During the training phase, developers who use copyrighted works without authorization as training data may infringe reproduction and adaptation rights. Although some assert the "non-expressive use" defense [6], fair use arguments become difficult to sustain when training data encompasses the core content of works and the generated results create market substitution effects. In the content generation phase, AI outputs that exhibit substantial similarity to protected works, such as mimicking a specific author's style or narrative framework, may trigger derivative copyright tort. Concurrently, significant disagreement exists regarding whether the originality of user prompts qualifies them as copyright subjects, with judicial practice tending to deny copyrightability to generated content lacking substantial human participation, resulting in a rights vacuum [7]. In the trademark domain, confusing use of commercial identifiers in AI-generated content, such as imitation of well-known brand visual elements, may constitute unfair competition, with courts increasingly restricting the defense space provided by the technology neutrality principle [8].

2.1.3 Indirect Infringement Scenarios

Beyond direct tort, Generative AI service providers may incur derivative liability through algorithmic manipulation and third-party misuse. Providers might manipulate output results through preset parameters—for example, directing users toward high-risk products in financial advisory contexts, thereby violating suitability obligations. Additionally, providers may bear algorithmic safety assessment responsibility when failing to filter contaminated information from training data, resulting in the generation of discriminatory content [9]. Regarding third-party misuse, developers may breach the principle of technological neutrality and be liable for contributory tort when they knowingly fail to implement effective content filters despite awareness of potential rights violations. A more profound systemic risk lies in algorithms amplifying social biases present in training data, such as gender discrimination in occupational image generation, leading to collective equal rights disputes. Such cumulative harms extend beyond the scope of traditional case-by-case remedies [10].

2.2 Core Characteristics of Tort Attribution

Generative artificial intelligence tort present four core characteristics at the attribution level. These characteristics directly constitute the structural origins of the application dilemma in traditional negligence liability, while also serving as the logical premise for strict liability to demonstrate its institutional advantages.

2.2.1 Technological Black Box and Causal Concealment

Tortious outcomes originate from algorithmic autonomous evolution mechanisms, with the inscrutability of deep neural networks leading to highly ambiguous technical causal chains between harmful conduct and resulting damages. Attribution of harm is not only difficult to anchor to specific technical nodes but is further complicated by model iterations occurring on weekly or daily cycles, creating an irretrievable discontinuity between the technical state at the time of the tortious act and the standards of subsequent judicial review [11]. This interweaving of dynamism and unobservability engenders fundamental cognitive barriers in the attribution process.

2.2.2 Systemic Risk Propagation

The homogeneous deployment of technical architectures enables single algorithmic flaws to transcend local constraints, instantaneously propagating throughout the entire domain via software update mechanisms, thereby transforming microscopic design errors into systemic vulnerabilities. Concurrently, influenced by network effects, the superlinear correlation between user base growth and risk intensity causes behavioral scale expansion to manifest as exponential amplification of social costs. The autonomous reinforcement of inherent data biases by algorithms further establishes self-perpetuating damage patterns, allowing infringement impacts to penetrate beyond individual boundaries into collective rights domains [12].

2.2.3 Dissolution of Responsibility Subjects within Network Structures

The multi-tiered division of labor across the industrial chain (fundamental research and development, data governance, system integration, scenario application) causes damage outcomes to be inherently embedded within technical collaboration networks, forming responsibility mappings characterized by multi-nodal behavioral coupling. The technological fluidity of open-source ecosystems intensifies the dissolution of entity boundaries, with the responsibility coordinates of original developers, commercial deployment entities, and end users continuously shifting alongside code modifications and scenario migrations. The stable binary "producer-user" relationship in traditional attribution frameworks is replaced by multidirectional responsibility flows. This diffuse nature results in preventive obligations being fragmentarily distributed throughout the entire technological chain, with interdependencies between various links generating self-cycling responsibility defense logic. Damage outcomes stem from the cumulative actions of multiple entities along the technological chain, with the absence of preventive measures at any single link potentially serving as grounds for attribution defense [13].

2.2.4 Irreversible Damage and High Remediation Costs

Once tort occurs, the consequences extend beyond controllable parameters, making complete elimination through ex post remedies exceedingly difficult. This manifests specifically in how false information generated by AI propagates at rates far exceeding traditional media, causing reputation restoration costs to increase exponentially; the substitution effect of generated content on original works' markets persists, with rights vacuums rendering losses difficult to quantify; algorithmic reinforcement of discriminatory patterns creates self-perpetuating cycles, requiring structural corrections that impose societal governance costs beyond individual cases [14]. These damage characteristics implicitly reveal a structural imbalance between remediation investments and damage scale, fundamentally constraining the efficacy of ex post remedies.

3. The Application Dilemma of the Negligence Liability Principle in Generative AI Infringement

Based on the aforementioned infringement characteristics, the traditional negligence liability principle encounters structural failure in the Generative AI domain, as its institutional logic struggles to adapt to technological realities. This section will thoroughly analyze the sources of this dysfunction, first revealing the judicial determination challenges of the "reasonable duty of care" standard under technological black boxes, dynamic iterations, and multi-party strategic interactions; second, demonstrating the institutional incentives that cause excessive expansion of "behavioral standards" and their amplified negative externality effects; finally, examining the inherent defects of alternative approaches such as presumption of negligence and product liability. Together, these three aspects indicate that the negligence liability framework faces insurmountable adaptability challenges in terms of attribution logic, behavioral incentives, and institutional costs, urgently requiring a paradigm shift.

3.1 Challenges in Judicial Determination of the Reasonable Standard of Care

In the context of Generative AI infringement scenarios, the core deficiency of the negligence liability principle is manifested in the ineffectiveness of judicial determination of the "reasonable care" standard. According to traditional tort law theory, the normative function of the negligence liability principle relies on judicial authorities establishing the actor's reasonable duty of care through case-by-case discretion. In law and economics, this process employs the Hand Formula as its fundamental analytical tool, requiring judges to engage in counter-factual analysis. If the cost of preventive measures not taken by the actor is lower than the expected accident loss (i.e., the product of the accident loss amount and its probability of occurrence), then the actor's behavior constitutes negligence [15]. Scholars generally consider the Hand Formula to be essentially a marginal efficiency test; when an actor fails to implement a certain marginal preventive measure—that is, when its cost is lower than the marginal safety benefit brought by the measure—negligence should be established [16]. Based on theoretical analysis using the Hand Formula, negligence liability can incentivize actors to adopt reasonable levels of care and select efficient preventive measures to avoid tort liability.

At the operational level, this powerful law and economics analytical paradigm requires plaintiffs to prove defendant negligence by identifying, discovering, and demonstrating effective preventive measures that defendants failed to implement. However, this framework faces structural failure under the combined effects of technological opacity and multi-party involvement. When damages occur, plaintiffs can leverage "hindsight" [17] to retrospectively assert "measures that should have been taken," such as demanding developers increase adversarial testing iterations during training or deploy real-time content filtering systems. Yet the rapid iteration of generative AI—evolving weekly or even daily—creates a temporal disconnect: safety technologies not yet commercialized when incidents occur may become industry standards during litigation proceedings, forcing judicial arbiters to establish "reasonable care" benchmarks in dynamic technological environments, with judgments often lagging behind technological realities. The more fundamental

challenge lies in the inherent delay in validating marginal preventive measures' effectiveness; while increasing Reinforcement Learning from Human Feedback iterations may reduce model hallucination probability [18], verifying actual effects requires massive data validation—an empirical technical burden litigation processes simply cannot accommodate.

This predicament is further amplified in scenarios involving multiple responsible entities. The generative artificial intelligence industry chain encompasses multiple stages including foundational model development, data cleansing, system deployment, and end-user applications, forming a network-like liability structure. Not only plaintiffs but also defendant tortfeasors will undoubtedly exert maximum effort to demonstrate that numerous other parties, including the plaintiff, could have implemented efficient preventive measures. They will argue that other entities could have adopted preventive measures with lower marginal costs and higher marginal safety outputs, thereby attempting to establish defenses based on contributory negligence or comparative negligence to absolve or at least mitigate their own liability [19]. In this context, the marginal analysis of the Hand Formula becomes distorted into a tool for deflecting responsibility, with defendants transforming litigation into an evidentiary competition to identify the "least-cost avoider" by arguing that other links in the chain presented lower-cost prevention opportunities. Ultimately, courts may resort to hasty mediated settlements after expending substantial resources evaluating the reasonableness of multiple parties' prevention costs, unable to verify any party's claims due to algorithmic opacity.

The algorithmic black box constitutes a deeper evidentiary paradox. The inexplicability of generative AI algorithmic decision-making processes traps the judicial system in a triple cognitive dilemma [20]. Plaintiffs struggle to pinpoint the technical root causes of harm—whether it stems from training data bias, prompt injection, or parameter overfitting. Meanwhile, defendants shield critical evidence behind trade secret defenses, forcing judges to make difficult determinations amid conflicting expert testimonies. This information asymmetry fosters severe adverse incentives, where developers may deliberately minimize security testing documentation to avoid creating potential liability evidence. Implementers tend to prefer closed-source models with reduced explainability to obscure responsibility boundaries through technical complexity. Consequently, the judiciary cannot establish stable reasonable care standards for adjudication nor develop effective behavioral guidance through case law, resulting in exponentially escalating institutional costs. When litigation costs approach or exceed damage compensation amounts, rational victims are compelled to abandon pursuit of accountability, resulting in substantial paralysis of the preventive function of negligence liability principles. The more profound harm lies in how the ambiguity of reasonable care standards distorts industrial behavior; enterprises may adopt excessive preventive strategies or conversely engage in risk speculation to avoid litigation risks. The structural disadvantage of judicial adjudication in technical cognition renders it both powerless to establish the baseline of "reasonable care" and incapable of penalizing "boundary-pushing" behaviors.

This demonstrates that even with quantitative analysis based on law and economics, determining a reasonable standard of care remains challenging. Theoretical consistency has not reduced operational disputes but instead incurred higher institutional costs. The primary difficulties in applying negligence liability to generative artificial intelligence torts lie in the complex causal mechanisms and the multiplicity of potentially negligent parties, which may lead to significant accountability uncertainty and impede victims' access to compensatory remedies. More critically, when the judicial determination costs for reasonable care in certain tort disputes consistently exceed their social benefits—that is, the total utility of preventing future accidents through adjudication — the negligence liability principle loses its foundational justification [21]. Generative AI infringements are approaching this critical threshold, as judicial assessment costs for reasonable care standards increase dramatically due to technological complexity, while the benefits of preventing future incidents diminish with rapid technological iteration. In this context, the judicial predicament regarding reasonable care standards is no longer a technical flaw but has become a central driver for paradigmatic liability reform.

3.2 Excessive Expansion of Activity Levels

In Generative AI infringement scenarios, the deficiency of negligence liability attribution lies not only in the unverifiability of the reasonable care standard but also in its systematic inflation of activity levels. In classical economic analysis models of tort liability, activity level specifically characterizes the frequency, scale, or intensity of risk-generating activities under given technological conditions, typically exemplified by vehicle mileage. In the generative AI domain, this manifests concretely in model deployment scope, user interaction frequency, and technological iteration speed. According to the economic analysis framework of tort law, negligence liability externalizes residual accident costs to society after reasonable care is taken, causing the marginal private cost curve of actors to fall significantly below the marginal social cost curve [22]. This institutional arrangement may distort the pricing mechanism of risk activities; when achieving reasonable prevention levels constitutes liability exemption, the marginal cost for actors to expand their operational scale approaches zero, while marginal private benefits remain fully internalized. This fracture in the cost-benefit structure inevitably drives actors beyond socially optimal activity levels, causing risk production scale to persistently deviate from Pareto efficiency boundaries.

The essence of behavioral level expansion is the institutionalized expression of negative externalities. In the field of generative AI, the theoretical core of the behavioral level has been reconstructed. First, technological endogeneity makes model iteration itself a core dimension of the behavioral level, with qualitative changes such as parameter growth and architectural updates directly amplifying risk exposure. Second, network effects create a superlinear relationship between user base and systemic risk [23], meaning behavioral level expansion no longer follows traditional laws of diminishing marginal returns. When negligence liability principles disconnect technological

evolution from risk assumption, companies naturally direct resources toward reducing marginal private costs rather than marginal social costs, which paradoxically leads to technological innovation bottlenecks—diluting unit liability costs through unlimited behavioral level expansion rather than achieving fundamental risk reduction through technological breakthroughs.

This institutional distortion will further induce a dual efficiency loss. Specifically, at the static level, capital allocation skews toward sectors with low liability sensitivity. Since residual risks under negligence liability remain unpriced, high-value but accident-prone research directions such as medical diagnostic algorithms face financing constraints. Conversely, low-risk, low-value applications like entertainment generation tools receive excessive investment due to reduced liability costs, creating resource misallocation that continuously expands potential social welfare losses. At the dynamic level, safety research and development investments become systematically suppressed. Actors' pursuit of liability exemption certainty leads them to strictly limit preventive investments to judicially recognizable "reasonable care" parameters while rejecting innovative safety solutions requiring long-term commitment. Technological evolution becomes confined within existing paradigms, creating a negative feedback loop of "stagnant care standards amid expanding behavioral scope." More problematically, the replicability of generative AI risks creates a multiplier effect for behavioral expansion [24]. In traditional infringement scenarios, single actions typically correspond to localized damages, but the homogeneous deployment of generative AI models means that individual design flaws can instantaneously propagate system-wide through software updates, transforming micro-level behavioral decisions into macro-level systemic risks. Negligence liability in this context completely loses its function of internalizing social costs, instead becoming an institutional engine for risk scaling.

3.3 Inherent Defects in Alternative Attribution Approaches

Regarding the dilemma of applying negligence liability to Generative AI infringement, presumption of negligence and product liability are viewed as potential alternative approaches [25]. However, their institutional frameworks remain constrained by the fundamental contradiction between technological opacity and multi-party attribution. These approaches not only fail to overcome existing limitations but also generate new systemic deficiencies during the process of liability transfer.

The application of the presumption of negligence principle in determining the liability for generative artificial intelligence infringement takes into account the aforementioned difficulties in establishing reasonable care. The logic behind presumption of negligence is that, although plaintiffs may struggle to provide direct evidence, accidents are likely caused by the actor's negligence; therefore, if defendants cannot prove they have fulfilled their duty of reasonable care, they will be presumed negligent. Intuitively, the presumption of negligence principle appears to mitigate issues of increased disputes and excessive costs resulting from pure negligence liability, thereby reducing the plaintiff's burden of proof. However, since any defendant presumed to be at fault still has the opportunity to rebut this presumption, the multi-party liability avoidance game persists, and the allocation of burden of proof may become even more complex. When a harmful incident occurs, defendants can still argue, based on the complexity of the technical chain, that the damage stems from marginal preventive measures not taken by third parties, or attribute it to defects in the victim's own conduct. This defense inevitably triggers a cost-benefit analysis of alternative preventive pathways, forcing the judicial system to return to the quantification dilemma of the Hand Formula. The more fundamental contradiction lies in the dynamic learning characteristics of generative artificial intelligence, which creates a paradox in demonstrating "reasonable care." Developers cannot fulfill their burden of proof if they refuse to disclose core algorithms, yet comprehensive disclosure of technical details would jeopardize trade secret protection [26]. The essence of this dilemma is the mapping of technological unverifiability into procedural rules, resulting in a presumptive mechanism that makes it difficult to achieve the initial goal of reducing litigation costs by circumventing the difficulty of determining the standard of reasonable care, and is instead preferable to the direct application of strict liability.

Product liability, as an alternative to negligence liability in the generative AI infringement discussion, centers on holding the designers, producers, and sellers of the generative AI product that caused the injury liable in tort for injuries caused by defects in that product, with the victim needing to prove that the product was defective in design or manufacture [27]. However, there are significant limitations to the scope of application of product liability in its current form in this area. The primary problem is the limitation of its “product” attribute. Generative AI infringements often arise from the provision of services or assisted decision-making processes, rather than physical “products”, which results in the exclusion of these situations from the scope of liability [28]. Secondly, although product liability is often regarded as strict liability, the very concept of “defect”, the central key to its determination, implies a reasonable care analysis. Defects usually mean that the safety of a product does not meet the reasonable expectations of consumers, or that its safety risk exceeds the utility it is intended to achieve, and imperfections in product design and production do not guarantee absolute safety, and do not necessarily constitute defects [29]. As a result, the determination of defects is still essentially an analytical process based on reasonable care or efficient prevention, which makes product liability still a de facto negligence liability in essence, and thus cannot really solve the problem of attribution of negligence liability in the field of Generative AI infringement and the corresponding system costs [30]. The deeper problem lies in the division of labor in the industry, which makes the advantage of product liability in limiting the scope of the responsible party a tool to get rid of responsibility. When an open source base model is deployed by a third party in a high-risk scenario, the original developer is removed from the liability system because he or she is not a producer. Even in a closed system, a misuse defense clause can lead to circularity - the essence of whether a user has violated the code of practice is still a matter of determining the level of reasonable care - and return product liability to the dilemma of applying the principle of attribution by negligence

described above.

4. Institutional Advantages of Applying Strict Liability Principles to Generative AI Infringement

In response to the systemic failure of negligence liability, strict liability principles demonstrate significant institutional advantages, providing a more suitable attribution pathway for generative AI infringement. This section focuses on the core advantages of strict liability principles. First, it demonstrates how simplifying attribution requirements effectively circumvents the challenges of technological black boxes and multi-party strategic interactions, significantly reducing institutional costs. Second, based on law and economics models, it systematically explains how strict liability simultaneously optimizes both the "level of care" and "activity level" of actors, guiding risk activities to spontaneously converge toward social optimality.

4.1 Endogenous Optimization of Due Care Standards

Compared to negligence liability, strict liability demonstrates profound institutional rationality in its impact on actors' due care levels. The core characteristic of strict liability lies in its ability to eliminate the judicial burden of retrospectively determining reasonable duty of care. Judges need not reconstruct the reasonableness of preventive measures within technical black boxes after harm occurs, nor must plaintiffs prove the defendant's fault. This transformation in attribution mechanisms directly resolves the application dilemma of the Hand Formula in generative artificial intelligence contexts — the proliferation of "hindsight bias" caused by rapid technological iteration and the responsibility-shifting game among multiple entities no longer constitute preliminary obstacles to liability determination.

It should be particularly clarified that strict liability, while significantly increasing the risk of damage compensation borne by actors, does not necessarily induce irrational behavior of excessive precaution. According to basic principles of law and economics, when the "reasonable duty of care" established by law precisely corresponds to the socially optimal level that minimizes the sum of accident costs and prevention costs, there is no difference between negligence liability and strict liability in incentivizing actors to adopt optimal precaution levels. The underlying mechanism is that under any liability attribution principle, rational actors aim to minimize their private costs. Even within a strict liability framework, where actors must bear all social costs including residual accident costs and prevention costs, they will still choose the level of precaution that minimizes their private costs—this level precisely coincides with the optimal prevention point that minimizes total social costs [31]. In other words, although strict liability mandates the internalization of residual risks, it does not alter the decision-making logic of prevention investment based on cost-benefit analysis by actors, and excessive precaution is not an inevitable consequence.

This theoretical derivation gains stronger practical vitality in the context of generative artificial intelligence infringement. Strict liability allocates residual risk to the tortfeasor, providing institutional space for autonomous optimization of preventive measures based on cost minimization objectives. Compared to the structural disadvantages of judicial authorities in technical cognition and ex post judgment, actors at the industry frontier possess significant professional advantages and dynamic adjustment capabilities. They can capture the security implications of algorithmic vulnerabilities in real-time and flexibly allocate R&D resources to balance risk control and innovation efficiency. The essence of this decentralized prevention decision mechanism is returning complex technology risk management authority to the actors with the greatest informational advantages and response capabilities, approaching socially optimal prevention levels through market-based interactions. Although in specific scenarios actors may not precisely achieve theoretical optimality, compared to judicially-driven negligence determination models, their prevention efficiency and accident cost control capabilities clearly better align with the dynamic requirements of technological innovation [32].

Certainly, the determination of responsible parties still requires legislative policy clarification. While the principle of strict liability circumvents judicial challenges in determining multi-party negligence, it does not fundamentally resolve the question of "who should bear responsibility." Legislators may, based on policy considerations of risk source and prevention efficacy, preemptively designate developers, manufacturers, and other key entities as liability bearers, without being constrained by complex post-event causal chain evidence. Such limitation of liability scope not only aligns with the normative goal of the least-cost avoider principle but also preserves sufficient negotiation space for market participants. Even if the initial liability allocation exceeds the actual optimal preventer, under controllable transaction costs, the Coase theorem suggests that relevant parties can still achieve liability redistribution through agreements and insurance mechanisms, ultimately directing risk toward the entity with the most efficient prevention capabilities [33].

4.2 Endogenous Regulation at the Behavioral Level

The principle of strict liability creates incentives for actors not only in terms of their level of care but also affects their behavioral level. Given a constant level of care, the choices made by actors regarding their behavioral level inherently influence the social costs of behavior, which constitutes the core difference between strict liability and negligence liability. The two liability regimes generate distinctly different incentive effects on actors' behavioral levels, directly addressing the systemic dilemma of negligence liability leading to excessive expansion of behavioral levels [34]. As previously discussed, a significant defect of the negligence liability principle is that once an actor meets the reasonable care standard, their marginal private cost for expanding activity levels approaches zero, while the residual accident costs are externalized to society. This creates excessive incentives for actors to continuously increase their activity levels to maximize private benefits, such as market share or data accumulation, resulting in risk production scales far exceeding socially optimal levels, ultimately institutionalizing negative externalities. The strict liability principle fundamentally corrects this distortion by requiring actors to bear responsibility for accident damages regardless

of whether they have met reasonable care standards. This implies that even after implementing optimal preventive measures, the actor must internalize residual accident costs through liability compensation. Consequently, when selecting a level of conduct, the actor must incorporate the expected accident costs, including residual accident costs, into their private cost function for consideration. A rational actor will choose a level of conduct that maximizes the net value of their activity—total benefits minus total private costs, which include both prevention costs and expected accident costs. This autonomously selected level of conduct by the actor precisely aligns with the optimal level of conduct required for social welfare maximization.

It is evident that the standard of conduct under strict liability is necessarily lower compared to negligence liability. However, this does not represent a loss of efficiency, but rather the inevitable result and efficiency manifestation of internalizing external costs. The value creation and risk generation in generative artificial intelligence activities often coexist and fluctuate in tandem. The optimal level of conduct is not a single-dimensional extreme point of either value or accident costs, but rather the peak point of net social benefit (total value minus total social costs, including prevention costs and accident costs). Strict liability guides actors to naturally converge to this efficiency frontier by compelling them to fully bear the social costs of their activities. In contrast, negligence liability, by severing the expansion of behavioral standards and the assumption of residual risks, inevitably drives systematic inflation of behavioral standards, leading to inefficiency. Therefore, arguments against strict liability based on "suppressing the scale of innovation activities" are difficult to sustain. The legitimacy of generative artificial intelligence research and application activities stems from creating positive social value, with risks being a byproduct. The lower behavioral standard guided by strict liability is essentially an efficiency level that eliminates external distortions, which is prudent and aligned with overall social interests. More importantly, this adjustment is the result of actors' autonomous decision-making based on cost-benefit analysis, possessing dynamic adaptability. Actors can reduce the residual accident costs per unit of activity through continuous technological innovation, thereby rationally elevating their activity levels under controlled risk conditions, achieving sustainable growth in net benefits. Conversely, under negligence liability, since actors need not bear residual accident costs, their intrinsic motivation to fundamentally reduce societal risks through technological innovation is weakened, making them more susceptible to falling into a vicious cycle of "stagnant care levels—expanding activity levels." [35]

Concurrently, to alleviate the irrational apprehension of liable entities regarding potential substantial compensation, particularly considering the possibility of inadequate insurance coverage, and to serve specific industrial innovation policy objectives, a liability cap mechanism can be incorporated within the strict liability framework. This rule restricts compensation amounts to statutory limits, and as long as this cap is below the actual damages incurred by the injured party, it can effectively enhance behavioral standards—as it reduces the expected per-incident costs for the actor [36]. Simultaneously, the certainty of compensation brought about by liability limits contributes to the formation and development of the liability insurance market. A mature insurance market not only further stabilizes behavioral expectations but also, through risk pooling mechanisms, more effectively provides remedies to victims on an average basis.

5. Conclusion

This thesis demonstrates through systematic analysis from a law and economics perspective that generative artificial intelligence presents a unique tort risk structure — characterized by technical opacity and causal concealment, systemic risk proliferation, network dissolution of liable entities, and irreversible harm—which leads to systematic failure of traditional negligence liability principles in terms of attribution logic, behavioral incentives, and institutional costs. The structural impediments identified through reasonable care not only result in prohibitive litigation costs and paralysis of preventive functions, but also, by severing behavioral expansion from residual risk allocation, intrinsically drive excessive proliferation of risk production, institutionalizing negative externalities. In contrast, strict liability principles, through fundamental simplification of attribution requirements, effectively circumvent technical opacity and multi-party game theory dilemmas, significantly reducing institutional operational costs and enhancing litigation efficiency. More critically, it optimizes both the attention level and behavioral level of actors through a mechanism of forced internalization of risk costs. On one hand, it incentivizes the most informationally advantaged parties to autonomously seek cost-effective preventive measures; on the other hand, it compels actors to incorporate all social costs of behavioral expansion into their considerations, thereby guiding risk activities to spontaneously converge toward socially optimal boundaries. This not only represents a fundamental correction to the distortion of the negligence liability system, but also constitutes an inevitable requirement for achieving dynamic equilibrium between risk and innovative activities. Therefore, under existing legal and technological conditions, shifting toward a strict liability principle is not only a more efficient attribution pathway for addressing the complexities of generative artificial intelligence infringement, but also a more feasible and necessary institutional choice that balances risk control, victim remediation, and the sustainability of technological innovation.

Acknowledgement

This paper is supported by the fund project: “Philosophy and Social Sciences Planning Project of Anhui Province” (AHSKQ2023D010)

References

[1] Laura Weidinger, et al., Taxonomy of Risks Posed by Language Models, Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), 214–229.

[2] Robert Cooter & Thomas Ulen, Law & Economic 6th.ed, Prentice Hall, 2011, p.211-213.

[3] Suzanne Beck, The Problem of Attributing Legal Responsibility in the Case of Robotics, 31 AI & Society 1, 1-9(2015).
[4] Huang Pei, The Challenge of Generative Artificial Intelligence to Personal Information Protection and Its Risk Regulation, 46 Modern Law Science 103, 103-106 (2024).
[5] Wang Liming, Legal Responses to Infringement in Generative Artificial Intelligence: Challenges and Solutions, 5 China Journal of Applied Jurisprudence 27, 27-38(2023).
[6] Zhang Yuanyuan, The Improvement of the Limitation of Copyright System of Machine Learning in the Era of Artificial Intelligence: An Addition to the Non-Expressive Use Theory, 46 Journal of Dalian University of Technology (Social Sciences) 92, 92-101 (2025).
[7] Jiang Ge, On the Copyrightability of AI-Generated Content: From the Perspective of User's Original Expression, 1 Intellectual Property 36, 36-67(2024).
[8] Li Zonghui, Legal Regulation of Artificial Intelligence Trademark Infringement and Unfair Competition: A Comprehensive Study, 1 China Trademark 29, 29-32 (2020).
[9] Omri Ben-Shahar, Data Pollution, 1 Journal of Legal Analysis 104, 104-159(2019).
[10] Cui Zijing, The Crisis and Response in the Protection of Equal Rights under Algorithm Discrimination Challenges, 3 Science of Law (Journal of Northwest University of Political Science and Law) 29, 29-42 (2019).
[11] Steven Shavell, Liability for Harm Versus Regulation of Safety, 2 Journal of Legal Studies 357, 357-374(1984).
[12] Danielle Keats Citron, Frank Pasquale, The Scored Society: Due Process for Automated Predictions, 1 Washington Law Review 34, 1-34(2014).
[13] Xie Yaowen, Matching Administrative Regulation and Tort Liability in the Generative AI Value Chain, 43 Tribune of Political Science and Law 36, 36-46(2025).
[14] Danielle Keats Citron, Technological Due Process, 6 Washington University Law Review 1249, 1249-1314 (2008).
[15] Richard A. Posner, A Theory of Negligence, 1 Journal of Legal Studies 29, 32-33(1972).
[16] Mark Grady, A New Positive Economic Theory of Negligence, 92 Yale Law Journal 799, 799-823(1983).
[17] Maggie Wittlin, Hindsight Evidence, 116 Columbia Law Review 1323, 1323-1394(2016).
[18] Matthew Dahl, et al., Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models, 16 Journal of Legal Analysis 64, 64-93(2024).
[19] Mark A. Lemley & Bryan Casey, Remedies for Robots, 86 The University of Chicago Law Review 1311, 1353(2019).
[20] Zhang Junhao, Research on Algorithmic Black Boxes: A Cognitive Science Perspective, 1 Studies in Science of Science 1, 1-14(2025).
[21] Richard Craswell & John E. Calfee, Deterrence and Uncertain Legal Standards, 2 Journal of Law, Economics and Organization 279, 279-303(1986).
[22] A. Mitchell Polinsky, An Introduction To Law and Economics, Wolters Kluwer, 2023, p.51-56.
[23] Guo Shuiwen, Xiao Wenjing, The Study on the Interaction Mechanisms of Network Effects, 4 Economic Review 14, 14-22(2022).
[24] Amos Tversky & Daniel Kahneman, Judgment under Uncertainty: Heuristics and Biases, 185 Science, New Series 1124, 1124-1131(1974).
[25] Deng Jianpeng & Zhao Songzhi, Regulation of Content-generated Artificial Intelligence: From the Perspective of Tort's Law, 3 Journal of Northeast Normal University (Philosophy and Social Sciences) 95, 95-104(2025).
[26] Liu Lin, Resolving the Conflict between Algorithm Interpretation Right and Trade Secret Protection, 2 Administrative Law Review 168, 168-176(2023).
[27] Chen Chen, The Subject of Liability and the Imputation Principles in Artificial Intelligence Tort, 37 Peking University Law Journal 363, 363-380(2025).
[28] Gabriel Weil, Tort Law as a Tool for Mitigating Catastrophic Risk from Artificial Intelligence, SSRN (Jan.28, 2024).
[29] Tao Liqin & Tong Jianhua, Legislative Reconstruction of the "Defect" Concept in the Product Quality Law of China, 2 Jianghan Tribune 120, 120-126(2025).
[30] David C. Vladeck, Machines Without Principals: Liability Rules and Artificial Intelligence, 89 Washington Law Review117, 117-150(2014).
[31] Steven Shavell, Foundations of Economic Analysis of Law, Belknap Press of Harvard University Press, 2004, p. 179-181.
[32] Hans-Bernd Schäfer & Claus Ott, The Economic Analysis of Civil Law, Edward Elgar, 2022, p. 217.
[33] George J. Stigler, The Theory of Price, Macmillan Co., 1966, p. 113.
[34] Steven Shavell, Strict Liability versus Negligence, 9 Journal of Legal Studies 1, 1-25(1980).
[35] Dai Xin, No-Fault Liability and Artificial Intelligence Development: A Perspective Based on Legal Economic Analysis, 5 ECUPL Journal 38, 38-55(2024).
[36] Hans-Bernd Schäfer & Claus Ott, The Economic Analysis of Civil Law, Edward Elgar, 2022, p. 219