# The Legal Duty of Safety-by-Design: Private Law Obligations of AI Developers Toward Youth Protection

Dr. Dilshad Ahmad Mhia-Alddin[1], and Dr. Ali Ehsan Teli [2]

[1st] International Relations & Diplomacy Department, College of Political Sciences, University of Duhok – Kurdistan Region – Iraq, dilshad.mhiaalddin@uod.ac

[2nd] Rand Private Institute, Duhok, Kurdistan Region – Iraq,ali.ehsan866@gmail.com

Corresponding Author: Dr. Dilshad Ahmad Mhia-Alddin*

## ABSTRACT

The fast growth of generative, interactive, and autonomous AI systems is creating new challenges for civil liability laws. Administrative rules and statutory oversight set basic market standards, but private law remains the primary means of corrective justice, risk management, and compensation. This study considers how 'Safety-by-Design' (SbD) has become a private-law duty for AI developers, notably regarding the protection of youth. For this study only, 'youth' and 'young people' refer to all persons below 18 years of age; persons aged 18 or older are outside its scope. This operational definition supports the paper's child-protection focus and does not replace statutory definitions or rule-specific age thresholds. It explores how engagement-focused algorithms, generative personas, and unchecked feedback loops can exploit youth's cognitive and neurological vulnerabilities. The analysis shows that these are deliberate engineering choices that fall under existing liability laws. By comparing civil and common law systems in places such as France, Germany, China, Iraq, the Kurdistan Region, the EU, the UK, and the US, this research examines how statutory protections and court decisions are converging. The findings show that digital companies can no longer shift the risks of algorithmic optimization onto vulnerable youth while escaping civil liability under current private law.

**KEY WORDS:** Safety-by-design; AI developer liability; private law obligations; youth protection; duty of care; product liability; algorithmic accountability; tort law and technology; AI governance regulation

## 1. INTRODUCTION

For this study, “youth” means all persons below 18 years of age, including younger children, and excludes persons aged 18 or older. This is an operational definition adopted for the paper’s child-protection analysis, not a claim that “youth” has a uniform international legal meaning. The United Nations uses ages 15–24 for statistical purposes, without prejudice to national definitions, and acknowledges that no universally agreed international definition of youth exists. By contrast, Article 1 of the Convention on the Rights of the Child (UN, 1990) defines a child as a person below 18, unless majority is attained earlier under the law applicable to that child. The study’s chosen boundary is also consistent with Article 106 of the Iraqi Civil Code No. 40 of 1951, which sets civil majority at 18 completed years; that provision defines majority rather than “youth.”

Within the analysis, “young people” and “children” refer to the same under-18 study population. References to “teenagers” mean persons aged 13–

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/**, **Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12597]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

17; references to “adolescents” concern a narrower developmental subgroup within the under-18 population and do not extend the study to adults. “Minors” is used descriptively for that population, but applicable law determines legal minority, capacity, consent, and entitlement to particular protections. The common research boundary does not imply that identical rules apply to every person below 18: each jurisdiction-specific rule must be assessed according to its own age threshold and other conditions. These terminology conventions do not alter the wording or scope of cited case titles, legal instruments, or source materials.

Base models, natural language processing, and persuasive algorithms currently shape the digital world. Software has changed from a simple data tool to an active, data-retrieval agent that can interact socially and influence how people behave. In these settings, harm to users usually comes from intentional design choices in neural network goals, behavioral reinforcement, and chat interfaces that seem human. Children and teenagers under 18 are especially at risk because their brains are still developing, they are more impressionable, and they often form strong attachments to these systems. This makes them more likely to be manipulated, experience negative reinforcement, or develop addictive behaviors.

This research is important because it reexamines how private law can regulate software design. Public regulations mostly focus on content moderation after the fact and broad protections for middlemen, but these are not enough for complex, autonomous AI systems. The study points out a difference between traditional private law, which deals with physical harm, tangible products, and adult users, and the psychological and developmental harm that minors face from poorly designed algorithms. It argues for clear, enforceable laws to stop companies from exploiting young people's vulnerabilities for profit.

This study uses a comparative, doctrinal, and practical approach. It reviews laws, legal scholarship, and recent court decisions from civil-law countries such as France, Germany, China, and Iraq/Kurdistan Region, as well as common-law countries such as the US and the UK, and EU-wide frameworks. The main idea is that Safety-by-Design is now a legal duty of care and a key factor in judging product defects. The paper is organized as follows: Chapter One looks at algorithmic harm; Chapter Two compares legal systems; Chapter Three discusses fault, design defects, and causation; Chapter Four reviews court trends and precedents; and Chapter Five presents practical steps and policy recommendations.

## 2. Chapter One: Conceptual Foundations and the Private Law Imperative of Safety-by-Design

### 2.1. Civil-Law Foundations of Preventive Care

Safety-by-Design should be situated within identifiable civil-liability provisions rather than asserted as a uniform duty arising whenever a developer creates a possible danger into the market. In Iraq, Article 202 of the Civil Code No. 40 of 1951 provides for compensation for harmful acts causing personal injury, while Article 204 provides the residual basis for compensation for wrongful infringement causing other damage. For an AI developer, these provisions provide potential grounds for liability where an unsafe design or omission is legally wrongful and causes compensable injury; introducing a risky system alone does not establish every element of liability. More specifically, Article 231 addresses mechanical machines and other things requiring special care to prevent harm: the person having them under their control is responsible for the damage they cause unless that person proves that sufficient precautions were taken to prevent it. This is the provision most directly relevant to the argument that control of a dangerous instrument requires preventive care.

Published Iraqi legal commentary has proposed applying Article 231 to AI systems; however, this is an interpretive position, not an express statutory classification of all AI software or a demonstrated AI-specific judicial holding. Its application therefore requires justification of both the system's classification as a qualifying thing and the defendant's relevant control over it.

The comparative statutory bases are likewise specific: Articles 1240 and 1241 of the French Code civil establish liability for damage caused by fault, negligence, or imprudence. These provisions can support an allegation that a developer negligently created or failed to reduce a foreseeable design risk, subject to proof of fault, damage, and causation. In Germany, § 823(1) BGB imposes compensation for intentional or negligent unlawful injury to specified protected interests, including life, body, and health; the judicially developed doctrine of Verkehrssicherungspflichten concretizes this provision by requiring those who create or maintain a source of danger to take necessary and reasonable protective precautions. This preventive duty does not require eliminating every conceivable risk. In China, Civil Code Article 1165 establishes fault-based tort liability, while Article 1167 expressly permits requests to cease infringement, remove a nuisance, or eliminate danger where a tortious act threatens another person's personal or property safety. These provisions supply distinct compensatory and preventive foundations; they should not be treated as interchangeable rules of automatic liability merely for placing AI on the market. (Sophie Williams, Jonas Schuett and Markus Anderljung, 2026)

### 2.2. Safety-by-Design and Developer Responsibility

In AI, developers build adaptive systems whose algorithms control content delivery, agent interactions, user behavior, and monetization. The researcher's position is that, where use by minors and associated harm are reasonably foreseeable, the applicable standard of preventive care should account for those users' characteristics rather than assume an exclusively adult audience. Age-assurance measures, proportionate session limits, crisis-response features, and clear disclosures of a system's non-human nature are potential safeguards to assess against the particular risk and applicable legal framework. The civil-code provisions identified above do not themselves expressly require every such measure for every AI system; their relevance lies in assessing wrongful conduct, adequate precautions, and causally connected injury in the circumstances. (Sophie Williams, Jonas Schuett and Markus Anderljung, 2026)

Private law provides compensation and remedies after harm occurs, working alongside public regulation. Administrative laws set general product standards, but civil liability assesses whether a company exercised sufficient care in each case. If an algorithm causes serious mental or physical harm to a minor, private law holds the developer responsible for harm caused by their design. Safety-by-Design links software architecture to legal responsibility, requiring safety to be built in from the start rather than added later.

### 2.3. Youth Susceptibility and the ELIZA Effect

Youth are especially vulnerable to algorithms because of the 'ELIZA effect,' where users see automated systems as conscious or empathetic. For isolated teenagers under 18, a chatbot that acts like a friend can become their main source of emotional support. Developers who release these systems without strong safety measures to prevent self-harm or worsen depression take advantage of this vulnerability to increase engagement. Legally, marketing manipulative systems without basic protections is a civil wrong because it creates an unreasonable risk of harm. Design safeguards have different but increasingly similar legal foundations in civil and common law systems internationally.


- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/**, **Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12597]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

## 3. Chapter Two: Comparative Private Law Frameworks

The legal obligation to implement child-centric Safety-by-Design safeguards finds diverse yet converging statutory foundations across international civil and common law jurisdictions.

### 3.1. The French Civil Framework

In French private law, the civil liability of artificial intelligence developers is governed the legal duty to include child-focused Safety-by-Design measures is supported by different but increasingly similar laws in civil and common law countries around the world. Any act by an AI developer that falls below the standard of a prudent and diligent professional constitutes an actionable civil fault (faute délictuelle). French jurisprudence has long recognized an evolving "obligation of safety" (obligation de sécurité), which imposes heightened affirmative duties on commercial operators who design and market complex technological mechanisms. Simultaneously, under Articles 1245 through 1245-17 of the Code civil (which transpose European product liability rules) (France Code civil, 1804), software and digital control systems are categorized as products subject to strict civil liability. Under Article 1245-3, product defectiveness is assessed against the public's legitimate safety expectations, taking into account the system's foreseeable presentation and planned use. Deploying an emotionally interactive or generative AI model accessible to children without incorporating structural safeguards constitutes a design defect (défaut de conception), thereby rendering the developer strictly liable for any resulting personal injury or medically recognized psychiatric harm, without requiring proof of subjective fault. (Karlekar, 2025)

### 3.2. The German Civil Framework

German private law establishes developer liability under the Bürgerliches Gesetzbuch (BGB) (Bürgerliches Gesetzbuch BGB, 1896) through the law of delict, alongside the strict liability framework of the Produkthaftungsgesetz (ProdHaftG). (Gesetz über die Haftung für fehlerhafte Produkte ProdHaftG, 1989) Under § 823 Abs. 1 BGB, liability results from the breach of Verkehrssicherungspflichten (traffic safety duties), which legally obligate any person or enterprise creating a source of danger in commercial circulation to implement all technically feasible and economically reasonable measures to protect third parties from harm. In the context of AI systems accessible to youth, these duties mandate the architectural integration of safety guardrails to prevent cognitive and developmental injuries. Furthermore, under § 823 Abs. 2 BGB, civil liability automatically attaches upon the breach of a statutory protective norm (Schutzgesetz). Developers who violate child protection statutes—such as the Jugendschutzgesetz (JuSchG) (Jugendschutzgesetz JuSchG, 2002), the Jugendmedienschutz-Staatsvertrag (JMStV) (Jugendmedienschutz-Staatsvertrag JMStV, 2002-2026), or the harmonized requirements of the European Union Artificial Intelligence Act (EU, Artificial Intelligence Act, 2024), are directly liable in tort for resulting damages. Concurrently, under § 1 ProdHaftG (Gesetz über die Haftung für fehlerhafte Produkte ProdHaftG, 1989), developers face strict product liability for design defects (Konstruktionsfehler) and warning failures (Instruktionsfehler) that impair the physical integrity or health of minor users.

### 3.3. The Chinese Civil Framework

The Civil Code of the People's Republic of China (Civil Code of the People's Republic of China, 2021) provides an integrated tort regime governing digital and algorithmic harms. Under Book VII

(Tort Liability), Article 1165 anchors fault-based liability, while Articles 1194 through 1197 govern the specific tortious obligations of network service providers and digital platform operators. Under these provisions, network entities that fail to carry out necessary, proactive technical measures to halt or prevent infringements of users' personal rights and mental well-being are held jointly and severally liable. This civil baseline is reinforced by Article 1202 of the Civil Code, which establishes strict product liability on manufacturers for design and manufacturing defects. When applied to AI systems, these articles operate in tandem with Chapter V ("Online Protection") of the PRC Law on the Protection of Minors (Law of the People's Republic of China on Protection of Minors, 2020) and the Provisions on the Management of Algorithmic Recommendations in Internet Information Services (Provisions on the Administration of Algorithmic Recommendation in Internet Information Services, 2022). These specialized instruments impose affirmative Safety-by-Design mandates: developers must establish youth-protection mechanisms, prohibit algorithmic push mechanisms that induce addiction or psychological distress, and institute automated screen-time and financial-expenditure caps. The failure to embed these architectural safeguards constitutes an actionable civil wrong under Article 1165, entitling injured minors and their guardians to compensatory and injunctive remedies.

### 3.4. The European Union Architecture

The European Union has established an integrated regional framework that synthesizes strict product liability, platform governance, and data protection to enforce Safety-by-Design obligations. The modern foundation of civil compensation is the Revised Product Liability Directive (EU, DIRECTIVE EU 2024/2853 on liability for defective products and repealing Council Directive 85/374/EEC, 2024), which replaces the 1985 regime. The Directive explicitly classifies standalone software, artificial intelligence systems, and digital manufacturing files as "products" subject to strict liability. Under Article 6, the evaluation of product defectiveness needs to account for the specific vulnerabilities of foreseeable user groups, such as children, as well as the model's capacity for post-deployment learning and autonomous drift. The Directive expands compensable damages to include medically recognized psychological harm and introduces procedural disclosure mechanisms and rebuttable presumptions of defectiveness and causation for claimants facing profound technical complexity.

Concurrently, Regulation (EU) 2024/1689 (the EU Artificial Intelligence Act) (EU, Artificial Intelligence Act, 2024) establishes prohibitive and mandatory architectural baselines. Article 5(1)(b) prohibits AI systems that exploit the vulnerabilities of specific groups because of age or mental disability to distort behavior in a manner that causes significant harm materially. For high-risk models, Chapter III mandates continuous risk management systems, technical robustness testing, and human monitoring mechanisms. Under the Digital Services Act (EU, The Digital Services Act, 2022), Articles 28, 34, and 35 require very large internet platforms to ensure high levels of privacy and safety for minors by default and eliminate algorithmic architectures that fuel systemic mental health harms. Finally, Article 25 of the General Data Protection Regulation (Regulation EU 2016/679) legally cements Data Protection by Design and by Default, prohibiting automated behavioral profiling and the non-consensual harvesting of minors' data.

### 3.5. The United Kingdom Framework

United Kingdom law addresses algorithmic Safety-by-Design through the interaction of common-law


- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12597]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

negligence, statutory product liability, and specialized child-safety legislation. (Kohl, 2024) Under English tort principles, rooted in Donoghue v. Stevenson [1932] AC 562 and refined in Robinson v. Chief Constable of West Yorkshire Police [2018] UKSC 4, an AI developer owes an affirmative duty of care to all foreseeable individuals when introducing a new source of physical or recognized psychiatric harm into the market. Deploying generative agents accessible to children without safeguards against self-harm inducement constitutes actionable negligence. Under Part I of the Consumer Protection Act 1987 (UK, Consumer Protection Act, 1987), strict liability attaches to defective software associated with consumer products. This private law baseline is reinforced by the Online Safety Act 2023 (UK, Online Safety Act, 2023), which, in Sections 11 and 12, imposes mandatory "Safety Duties for Children" on regulated online services, requiring them to use Safety-by-Design to prevent minors from encountering priority harms, such as suicide and self-harm material. Furthermore, Section 123 of the Data Protection Act 2018 (UK, Data Protection Act, 2018) establishes the statutory Age-Appropriate Design Code (Children's Code), requiring developers to treat the best interests of the child as a primary consideration by turning off behavioral profiling and manipulative "nudge techniques" by default.

### 3.6. The United States Framework

In the United States, civil claims against AI developers operate under the framework of strict products liability and common-law negligence. Under the Restatement (Third) of Torts: Products Liability § 2(b) (Foundations of Business Law and the Legal Environment, 1997), a product is defective in design when the foreseeable risks of harm could have been reduced or avoided by the adoption of a Reasonable Alternative Design (RAD), and the omission of that alternative renders the product not reasonably safe. In youth safety litigation, plaintiffs argue that algorithmic recommendation engines, infinite scroll interfaces, and unconstrained conversational personas constitute defectively designed software products. Developers face strict liability because feasible alternative designs—such as deterministic content filters, verified age-gating, and automatic session limiters—were readily available but bypassed to maximize platform engagement. Concurrently, federal courts are dismantling historic defenses under Section 230(c)(1) of the Communications Decency Act of 1996 (47 U.S.C. § 230) (SECTION 230 OF THE COMMUNICATIONS DECENCY ACT OF 1996), establishing that proprietary algorithmic curation and generative outputs represent the developer's own expressive conduct and product architecture rather than protected third-party speech. At the state level, statutory enactments such as the California Age-Appropriate Design Code Act (the California Age-Appropriate Design Code Act, 2023) enforce mandatory Data Protection Impact Assessments and child-protective default configurations.

### 3.7. The Iraqi and Kurdistan Regional Civil Framework

Under Iraqi private law, codified in the Iraqi Civil Code No. 40 of 1951 (the Iraqi Civil Code, 1951) and enforced across both federal Iraq and the Kurdistan Region, developer liability is grounded in the general theory of harmful acts and specialized provisions governing dangerous objects. Article 204 of the Iraqi Civil Code establishes the basic principle of extra-contractual liability: every that causes damage to another obligates the actor to make full reparation. In Iraqi civil doctrine, unlawful act or fault encompasses not only positive intentional acts but also negligent omissions, failure to uphold professional vigilance, and the distribution of unsafe mechanisms. Importantly, Article 205 of the Iraqi Civil Code explicitly

provides a statutory right to compensation for moral and psychological harm, granting courts the authority to award damages when an algorithmic system inflicts severe emotional distress, cognitive manipulation, or psychiatric injury on a minor.

Furthermore, Article 231 of the Iraqi Civil Code imposes a specialized presumption of liability on anyone exercising custody or control over machines or items that require special care to prevent harm. In contemporary doctrine, complex autonomous algorithmic architectures and generative AI models constitute sophisticated non-human instruments that require strict technical containment. Under Article 231, the developer or enterprise operating the AI system is held strictly liable for any resulting damages unless it proves that the harm arose from an external, unforeseeable cause beyond its control. This private law framework is reinforced by Iraqi Consumer Protection Law No. 1 of 2010 (Iraqi Consumer Protection Law, 2010), specifically Articles 6 and 7, which mandate that all commercial goods and digital services meet rigid safety specifications, remain free from hidden defects, and do not endanger the health or safety of consumers, thereby invalidating standard contractual disclaimers.

## 4. Chapter Three: Algorithmic Fault, Design Defects, and Causation in AI Systems

Applying private law to artificial intelligence systems requires re-evaluating foundational legal doctrines concerning breach of duty, design defectiveness, and proximate causation.

### 4.1. Fault-Based Liability and Foreseeability

In fault-based liability regimes, the developer's breach commonly takes the form of culpa in omittendo—the deli Applying private law to AI systems means rethinking basic legal ideas about breach of duty, design defects, and what counts as a direct cause of harm. If a system generates harmful, eating-disorder, or psychologically exploitative content and deploys it without deterministic containment filters, the legal test of foreseeability is fulfilled. Behavioral science and computational research show that generative conversational agents predictably induce the "ELIZA effect" (Karlekar, 2025), leading children to attribute real human consciousness, emotional empathy, and infallibility to automated outputs. The failure to foresee and address this predictable psychological effect constitutes actionable professional negligence.

### 4.2. Design Defects and Failures to Warn

Under strict product liability doctrines, the inquiry centers on the product's objective condition rather than the developer's subjective blameworthiness. An AI system is defective in design (Konstruktionsfehler in German law, défaut de conception in French law) when its internal algorithmic mechanics create unreasonable safety hazards that a safer alternative design could have eliminated. If an AI enterprise could have incorporated robust age-assurance, content-suppression algorithms, and automatic crisis-intervention routing without destroying the application's utility, omitting these measures renders the system legally defective. Concurrently, an instruction defect (Instruktionsfehler / défaut d'information) occurs when developers market an anthropomorphic chatbot to youth without supplying clear, comprehensible warnings to parents and minors about the agent's non-human nature and its psychological risks.

### 4.3. State-of-the-Art and Development-Risk Defenses

Furthermore, AI developers cannot shelter behind the "state-of-the-art" or development risk defense (risque de développement / Entwicklungsrisiko)


- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12597]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

(Meenu Mary John, Helena Holmstr¨om Olsson, and Jan Bosch, 2024) when the cognitive vulnerabilities of minors and the technical modalities for mitigating generative toxicity were widely documented in academic literature and industry best practices before commercial deployment. Where technical containment methods—such as reinforcement learning from human feedback (RLHF), safety classifiers, and rule-based guardrails—are part of the state of the art, a developer's failure to implement them precludes any statutory liability exemption.

### 4.4. Causation and Intervening Factors

The inquiry into causation in AI-induced harm requires establishing both factual causation (conditio sine qua non) and proximate causation (adäquater Kausalzusammenhang / lien de causalité direct et certain). Defense assertions that an autonomous AI generates emergent, non-deterministic outputs—or that third-party user engagements and the minor's pre-existing psychological vulnerabilities constitute intervening superseding causes (force majeure or fait de la victime)—fail under rigorous private law scrutiny. Where an algorithmic system is intentionally built to optimize user involvement through dynamic feedback loops that flexibly reinforce a minor's psychological vulnerabilities, the developer's failure to program deterministic protective constraints remains the legally operative and adequate cause of the resulting developmental injury. (Johann Laux and Hannah Ruschemeier, 2025)

## 5. Chapter Four: Emerging Jurisprudence and Court Precedents

Judicial bodies across major jurisdictions have increasingly imposed private-law liability on software and AI developers, progressively dismantling procedural and statutory immunities when design defects harm minor users.

### 5.1. United States Jurisprudence

In Garcia v. Character Technologies, Inc. (Garcia v. Character Technologies, 2025), a civil action was initiated following the suicide of a 14-year-old adolescent who had formed an intense, emotionally dependent parasocial relationship with an anthropomorphic AI conversational agent. The complaint asserted claims sounding in strict products liability for design defect, negligent design, and failure to warn. The federal court rejected the developer's assertion of absolute First Amendment protection and statutory immunity under Section 230 of the Communications Decency Act. The court held that where an AI platform is intentionally engineered to emulate human empathy, systematically exploits adolescent neurological vulnerabilities to maximize user retention, and fails to implement elementary age-verification or crisis-intervention guardrails, the action challenges the product's functional design architecture rather than third-party speech.

This holding built upon the Third Circuit's seminal decision in Anderson v. TikTok, Inc. (Anderson v. TikTok Inc, 2024), where the court reversed the dismissal of a wrongful death suit arising from a 10-year-old child's death resulting from a dangerous viral challenge delivered via her algorithmic recommendation feed. The Third Circuit determined that a platform's proprietary recommendation algorithm constitutes the enterprise's own first-party expressive conduct and product design, rendering Section 230 immunity inapplicable to claims targeting the algorithmic delivery mechanism.

These principles were reinforced in IN-Re-Social Media Adolescent Addiction/Personal Injury Products Liability Litigation (2023-2026), where Judge Yvonne Gonzalez Rogers held that tort claims directed at defective product functionalities—such as infinite scroll feeds, intermittent reward schedules, and deficient

parental oversight tools—challenge the engineering of the software rather than published speech. This trajectory is consistent with the earlier precedent in Lemmon v. Snap, Inc. (Lemmon v. Snap, 2021), which affirmed that software developers owe a traditional duty of reasonable care in product design to avoid creating features that foreseeably induce dangerous behavior among minor users.

### 5.2. France and European Jurisprudential Trends

In European civil practice, the Court of Justice of the European Union CJEU (CJEU, n.d.) and the French Cour de cassation (cassation, n.d.) has increasingly affirmed that the strict "obligation of safety" (obligation de sécurité) applies comprehensively to automated digital services and software environments. National courts applying the principles codified in the Revised Product Liability Directive (Directive EU, 2024/2853) have established that autonomous algorithmic drift, model intricacy, and opacity do not break the causal link between a developer's architectural omission and the resulting psychological damage suffered by a user. European courts have held that an enterprise that commercializes an algorithmic product bears an ongoing duty to monitor and maintain the system's protection and integrity throughout its entire lifecycle.

### 5.3. German Jurisprudence

In Germany, the Federal Court of Justice (Bundesgerichtshof - BGH) has consistently applied the doctrine of Verkehrssicherungspflichten (traffic safety duties) to commercial operators who introduce technological hazards into the marketplace. In recent civil determinations addressing automated and digital systems, the BGH has affirmed that enterprises that control algorithmic platforms have to establish organizational and technical safeguards that are proportional to the vulnerability of the protected group. (Beltran, 2025) Under § 823 Abs. 2 BGB, German courts treat violations of youth protection regulations and EU safety mandates as establishing a direct presumption of civil fault, enabling injured minors to recover full compensation for health impairments caused by manipulative digital designs.

### 5.4. Chinese Internet Court Precedents

In China, specialized judicial bodies, including the Beijing Internet Court (Court T. B., n.d.) and the Hangzhou Internet Court (Court T. H., n.d.), have delivered groundbreaking rulings enforcing platform tort liability under Civil Code Article 1165 and the Minor Protection Law (iapp, 2026). In actions brought against internet platforms regarding algorithmic addiction and exposure to damaging content, the Internet Courts have ruled that platform operators who design recommendation algorithms that push harmful, addictive content to minors fail to fulfill their statutory duty of care. The courts have issued mandatory civil injunctions requiring platforms to re-engineer their recommendation algorithms, enforce strict youth modes, and pay civil damages for moral and psychological harm inflicted upon minor users.

### 5.5. Iraqi and Kurdistan Regional Jurisprudence

In Iraqi and Kurdistan Regional civil practice, the Federal Court of Cassation has maintained a consistent jurisprudential doctrine governing tortious fault under Article 204 and custodian liability under Article 231 of the Iraqi Civil Code. In recurring rulings concerning complex electronic and mechanical instruments, the Court of Cassation has established that commercial manufacturers and operators are held to the rigorous standard of a specialized professional. The failure to provide


- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12597]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

necessary protective interlocks, fail-safes, or explicit operational warnings constitutes professional fault, triggering full compensatory obligations.

Furthermore, in applying Article 205 of the Civil Code, the Court of Cassation has repeatedly affirmed that substantial monetary damages must be awarded for moral and psychological harm suffered by minors and their families as a result of negligent commercial distribution. Applying this doctrine to modern digital and algorithmic systems, Iraqi and Kurdistan jurisprudence determines that an AI developer or commercial enterprise that deploys an uncontained, psychologically harmful generative model into the local online marketplace without safety-by-design mechanisms is fully liable for all resulting material, physical, and psychological injuries under Articles 204, 205, and 231 of the Iraqi Civil Code.

## 6. Conclusion

To meet the legal standard of reasonable care and fix design flaws, AI developers must translate general legal duties into clear, practical steps at every stage of the machine learning process.

During the initial training and fine-tuning phases, developers should execute targeted data filtering to remove training data containing self-harm instructions, severe harassment, eating-disorder promotion, and non-consensual exploitative material. Optimization objective functions should be separated from engagement-maximization metrics, such as session duration or user interaction rate, and instead aligned with developmental protection constraints. Developers must also document model training procedures, risk assessments, and red-teaming evaluations to demonstrate technical diligence and reduce civil exposure under product liability doctrines. At the inference and user interface stages, Safety-by-Design mandates the implementation of automated sentiment and crisis detection classifiers capable of identifying acute emotional distress or self-harm ideation in real time. When these classifiers activate, the AI system should terminate the conversation, prevent further harmful reinforcement, and immediately present localized crisis intervention resources. Additionally, systems that target or are accessible to youth must include privacy-preserving age-assurance mechanisms, enforce automatic session limits, disable algorithmic profiling by default, and provide transparent, auditable controls for parental and guardian oversight.

Modern civil liability shows that 'Safety-by-Design' is no longer optional or just a compliance formality. It is now a binding legal obligation grounded in the core principles of tort and product liability law. In many legal systems, including French, German, Chinese, Iraqi, and Anglo-American, courts and laws require AI developers to meet a clear standard of proactive care.

This research leads to three main legal conclusions. First, AI developers' duty of care cannot be delegated and must be based on what is reasonable for users under 18. Second, designs that encourage addictive use, emotional manipulation, or uncontrolled generative outputs are legally defective, removing traditional legal protections for developers. Third, current private law treats psychological breakdowns and diagnosed mental health harm as real injuries, allowing victims of algorithmic negligence to get compensation and court orders for relief.

### 6.1. Recommendations

To protect children and clarify AI law, this research offers multiple practical recommendations. Courts should continue to treat algorithmic design and content curation as the developer's own product, not as protected third-party speech, so that legal review is possible. Lawmakers in civil-law countries, including Iraq and the Kurdistan Region, should

officially classify software and AI models as 'products' under consumer and civil liability laws, making it easier to bring claims if algorithms are complex. Finally, AI developers should include Safety-by-Design features such as mandatory age checks, crisis-intervention tools, and robust parental controls to ensure innovation does not harm children's mental, emotional, or physical health.

## 7. REFERENCES


Anderson v. TikTok Inc, 22-3061 (U.S. Court of Appeals for the Third Circuit August 27, 2024). Retrieved from https://law.justia.com/cases/federal/appellate-courts/ca3/22-3061/22-3061-2024-08-27.html

Beltran, M. (2025). AI algorithms under scrutiny: GDPR, DSA, AI Act and CRA as pillars for algorithmic security and privacy in the European Union. Computers & Security, 158.

(1896). Bürgerliches Gesetzbuch BGB. Retrieved from https://shorturl.at/aibDh

cassation, L. C. (n.d.). Retrieved from https://www.courdecassation.fr/en

(2021). Civil Code of the People's Republic of China. https://shorturl.at/3SsSu.

CJEU, t. C. (n.d.). Retrieved from https://eur-lex.europa.eu/collection/eu-law/eu-case-law.html

Court, T. B. (n.d.). Retrieved from https://english.bjinternetcourt.gov.cn/

Court, T. H. (n.d.). Retrieved from https://www.ehangzhou.gov.cn/2025-07/31/c_294458.htm

EU. (2022). The Digital Services Act. Retrieved from https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng

EU. (2024). Artificial Intelligence Act. Retrieved from https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng/

EU. (2024). DIRECTIVE EU 2024/2853 on liability for defective products and repealing Council Directive 85/374/EEC. Retrieved from https://eur-lex.europa.eu/eli/dir/2024/2853/oj/eng

Foundations of Business Law and the Legal Environment. (1997). Retrieved from https://biz.libretexts.org/Bookshelves/Civil_Law/Foundations_of_Business_Law_and_the_Legal_Environment

(1804). France Code civil. Légifrance, https://shorturl.at/K5gNH.

Garcia v. Character Technologies, 6:2024cv01903 (U.S. District Court for the Middle District of Florida October 31, 2025). Retrieved from https://law.justia.com/cases/federal/district-courts/florida/flmdce/6:2024cv01903/433581/221/

(1989). Gesetz über die Haftung für fehlerhafte Produkte ProdHaftG. https://shorturl.at/BmgKb.

iapp. (2026). China sets a judicial benchmark for AI disputes. Retrieved from https://iapp.org/news/a/china-sets-a-judicial-benchmark-for-ai-disputes

IN RE: SOCIAL MEDIA ADOLESCENT ADDICTION/PERSONAL INJURY PRODUCTS LIABILITY LITIGATION, 4:22-md-03047 (District Court, N.D. California 2023-2026). Retrieved from https://www.courtlistener.com/docket/65407433/in-re-social-media-adolescent-addictionpersonal-injury-products-liability/

(2010). Iraqi Consumer Protection Law. Retrieved from https://shorturl.at/eNNz2

Johann Laux and Hannah Ruschemeier. (2025). Automation Bias in the AI Act: On the Legal Implications of Attempting to De-Bias Human Oversight of AI. European Journal of Risk Regulation , Volume 16, Special Issue 4: Special Issue on Climate Interventions Governance, pp. 1519 - 1534.

(2002-2026). Jugendmedienschutz-Staatsvertrag JMStV. https://shorturl.at/3uJ80.

(2002). Jugendschutzgesetz JuSchG. https://shorturl.at/efbyd.

Karlekar, S. (2025). The Psychology of the Eliza Effect: Anthropomorphism, Social Presence, and Projection in Human-AI Interaction. International Journal for Multidisciplinary Research (IJFMR), 7(6).

Kohl, U. (2024). Toxic recommender algorithms: immunities, liabilities and the regulated self-regulation of the Digital Services Act and the Online Safety Act. Journal of Media Law, Volume 16, Issue 2: Special edition: The Online Safety Act 2023, 1–35.

(2020). Law of the People's Republic of China on Protection of Minors. https://shorturl.at/y697I.

Lemmon v. Snap, 20-55295 (U.S. Court of Appeals for the Ninth Circuit May 4, 2021). Retrieved from https://law.justia.com/cases/federal/appellate-courts/ca9/20-55295/20-55295-2021-05-04.html

Meenu Mary John, Helena Holmstr¨om Olsson, and Jan Bosch. (2024). Artificial Intelligence in manufacturing: State of the art, perspectives, and future directions. CIRP Annals, 73(2).

(2022). Provisions on the Administration of Algorithmic Recommendation in Internet



- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/**, **Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12597]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

(1951). the Iraqi Civil Code. https://shorturl.at/1XwK6.
UK. (1987). Consumer Protection Act. Retrieved from https://www.legislation.gov.uk/ukpga/1987/43/contents
UK. (2018). Data Protection Act. Retrieved from https://www.legislation.gov.uk/ukpga/2018/12/contents
UK. (2023). Online Safety Act. Retrieved from https://www.legislation.gov.uk/ukpga/2023/50/contents
UN. (1990). Convention on the Rights of the Child. Retrieved from https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child
Information Services. Retrieved from https://shorturl.at/AteBn
SECTION 230 OF THE COMMUNICATIONS DECENCY ACT OF 1996. (n.d.). Retrieved from https://www.justice.gov/archives/ag/department-justice-s-review-section-230-communications-decency-act-1996
Sophie Williams, Jonas Schuett and Markus Anderljung. (2026). On Regulating Downstream AI Developers. European Journal of Risk Regulation, 94 - 122.
the California Age-Appropriate Design Code Act. (2023). Retrieved from https://california.public.law/codes/civil_code_section_1798.99.28

## پوختە

گەشەکرنا لەز یا سیستەمێن ژیرییا دەستکرد (AI) بێن بەرهەمهێنەر، کارلێککار، و خۆسەر، ئاستەنگێن نووی بۆ یاسایێن بەرپرسیارەتیا شارستانی دروست دکەت. ڕێسایێن کارگێڕی و چاڤدێریا یاسادانانێ پێوەرێن سەرەکی بێن بازاڕی دادەزرینن، لێ یاسایا تایبەت وەک ئامرازێ سەرەکی دمینیت بۆ دەستەبەرکرنا دادپەروەریا چاکسازیێ (ڕاستڤەکرنێ)، بڕێڤەبرنا مەترسیان، و دابینکرنا قەرەبوویێ. ئەڤ ڤەکۆلینە دیڤچوونێ د وێ چەندێ دا دکەت کا چەوا "پاراستن ژ دەستپێکا دیزاینێ (Safety-by-Design - SbD) "بوویە ئەرکەکێ یاسایا تایبەت ل سەر ملێن گەشەپێدەرێن ژیرییا دەستکرد، ب تایبەتی دەربارەی پاراستنا گەنجان. ب تنێ بۆ ڤێ ڤەکۆلینێ، دەستەواژەیێن "گەنج" و "کەسێن گەنج" ئاماژەیێ ددەنە هەمی وان کەسێن کو ژیێ وان ل ژێر ١٨ ساڵیێ یە؛ و کەسێن ژیێ وان ١٨ ساڵ یان پتر ژ دەرڤەی چوارچۆڤێ ڤێ ڤەکۆلینێ نە. ئەڤ پێناسەکرنا کارپێکری پشتەڤانیێ ل سەرەنجا ڤەکۆلینێ یا ل سەر پاراستنا زارۆکان دکەت و نەبوویە جهگۆڕا پێناسەیێن یاسایی یان ئاستێن دیارکری بێن ژی ل دویڤ ڕێسایان. ئەڤ ڤەکۆلینە دکۆلیتە د وێ چەندێ دا کا چەوا ئەو ئەلگۆریتمێن گرنگیێ ددەنە ڕاکێشانا سەرنجێ، کەسایەتیێن بەرهەمهێنەر، و بازنەیێن کارڤەدانێ بێن کۆنترۆلنەکری، دشێن مفا ژ لاوازیێن مەعریفی و دەماری بێن گەنجان وەربگرن و وان بقۆزنەوە. شیکاری دیار دکەت کو ئەڤە هەلبژارتنێن ئەندازیاری بێن ئەنقەست و ب مەبەستن کو دکەڤنە بن چوارچۆڤێ یاسایێن نوکە بێن بەرپرسیارەتیێ دا. ب ڕێکا بەراوردکرنا سیستەمێن یاسایا شارستانی و یاسایا گشتی (Common Law)ل دەڤەرێن وەکی فڕەنسا، ئەڵمانیا، چین، عیراق، هەرێما کوردستانێ، ئێکەتیا ئەورووپا، بەریتانیا، و ویلایەتێن ئێکگرتی بێن ئەمریکا، ئەڤ ڤەکۆلینە دیڤچوونێ د وێ چەندێ دا دکەت کا چەوا پاراستنێن یاسادانانێ و بڕیارێن دادگەهان بەرەڤ لێکنێزیکبوونێ دچن. ئەنجام دیار دکەن کو کۆمپانیێن دیجیتاڵی نێدی نەشێن مەترسیێن چاکسازی و گونجاندنا ئەلگۆریتمی بێخنە سەر ملێن گەنجێن لاواز ب مەبەستا خۆدزین ژ بەرپرسیارەتیا شارستانی ل ژێر یاسایا تایبەت یا نوکە.

**پەیڤێن سەرەکی:** پاراستن ژ دەستپێکا دیزاینێ، بەرپرسیارەتیا گەشەپێدەرێ ژیرییا دەستکرد، پێگری و ئەرکێن یاسایا تایبەت، پاراستنا گەنجان، ئەرکێ چاڤدێریێ و هشیاریێ، بەرپرسیارەتیا بەرهەمی، لێپرسینەوەیا ئەلگۆریتمی، یاسایا بەرپرسیارەتیا مەدەنی و تەکنەلۆجیا، ڕێکخستن و حوکمڕانیا ژیرییا دەستکرد.

## الملخص

يفرض النمو المتسارع لأنظمة الذكاء الاصطناعي التوليدية والتفاعلية والمستقلة تحدياتٍ جديدةً أمام قوانين المسؤولية المدنية. ففي حين تضع القواعد الإدارية والرقابة التشريعية معايير أساسية للسوق، يظل القانون الخاص الأداةَ الرئيسية لتحقيق العدالة التصحيحية، وإدارة المخاطر، والتعويض. تبحث هذه الدراسة في كيفية تحول مفهوم "السلامة بالتصميم" (Safety-by-Design) إلى التزامٍ قانوني خاص يقع على عاتق مطوري الذكاء الاصطناعي، لا سيما فيما يتعلق بحماية الشباب. ولأغراض هذه الدراسة، يُقصد بمصطلحي "الشباب" أو "صغار السن" جميع الأشخاص الذين تقل أعمارهم عن 18 عاماً؛ أما من بلغوا سن الثامنة عشرة فما فوق، فهم خارج نطاق هذا البحث. ويخدم هذا التعريف الإجرائي تركيزَ الدراسة على حماية الأطفال، دون أن يحل محل التعريفات التشريعية أو الحدود العمرية المحددة في قواعد قانونية أخرى. كما تستكشف الدراسة كيف يمكن للخوارزميات المصممة لتعزيز التفاعل، والشخصيات التوليدية (generative personas)، وحلقات التغذية الراجعة غير المنضبطة، أن تستغل نقاط الضعف الإدراكية والعصبية لدى الشباب. ويُظهر التحليل أن هذه الممارسات ناتجة عن خيارات هندسية متعمدة تندرج تحت طائلة قوانين المسؤولية القائمة. ومن خلال مقارنة أنظمة القانون المدني والقانون العام في دول ومناطق مثل فرنسا، وألمانيا، والصين، والعراق، وإقليم كردستان، والاتحاد الأوروبي، والمملكة المتحدة، والولايات المتحدة، يبحث هذا العمل في تقارب الحمايات التشريعية مع الأحكام القضائية. وتخلص الدراسة إلى أن الشركات الرقمية لم يعد بإمكانها تحميل الشباب -باعتبارهم فئةً ضعيفة- مخاطرَ التحسين الخوارزمي، مع الإفلات في الوقت ذاته من المسؤولية المدنية بموجب القانون الخاص الساري.


**الكلمات المفتاحية:** السلامة بالتصميم، مسؤولية مطوري الذكاء الاصطناعي، التزامات القانون الخاص، حماية الشباب، الالتزام بالرعاية، مسؤولية المنتج، المساءلة الخوارزمية، أحكام المسؤولية التقصيرية والتكنولوجيا، تنظيم حوكمة الذكاء الاصطناعي.


- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/**, **Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12597]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]