German Law Journal (2025), **26**, pp. 1404–1422
doi:10.1017/glj.2026.10198

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Reconstructing Liability Mechanism for Humanoid Robot Accidents

Weiwei Shen

China University of Political Science and Law, China
Email: shenww2036@163.com

(Received 30 May 2025; accepted 10 July 2025)

Abstract
Humanoid robots, with their capacity for human-like interaction and autonomous decision-making, present novel legal challenges in accident scenarios. This Article argues that existing Chinese accident law cannot fully accommodate accidents involving humanoid robots because hybrid human-algorithm control, adaptive machine learning, and embodied human-like interaction together destabilize traditional assumptions about fault, causation, proof, and remedy. To address these shortcomings, the Article proposes a reconstructed liability system that (1) establishes robust evidentiary processes for determining fault, (2) implements "reasonable person" technical standards for humanoid robot behavior, (3) designates manufacturers as the "least cost avoiders," and (4) considers behavioral correction and retribution for robots to address victims' psychological needs. This approach aims to foster an "accidental utopia" where innovation and safety are harmonized.

Keywords: Humanoid robots; machine learning; hybrid control; liability system

A. Introduction

Throughout history, as technological advancements have reshaped society, legal systems, particularly the law of accidents,[1] have been forced to evolve in order to address the new challenges these innovations introduce. From the invention of the automobile to the development of medical devices, the law has adapted to ensure that individuals and entities are held accountable for the harms caused by emerging technologies. These adaptations often involve reevaluating existing liability mechanisms and introducing new legal apparatuses to better fit the unique circumstances presented by novel technologies.

Humanoid robots represent the latest frontier in this ongoing process of legal adaptation. Designed to mimic human behavior, appearance, and even cognitive functions, humanoid robots are increasingly integrated into sectors ranging from healthcare and customer service to manufacturing and home assistance. Their potential to improve productivity, enhance elderly care, and create unprecedented opportunities is immense. However, they also introduce new categories of risk, as injuries caused by humanoid robots present unique challenges that current legal frameworks are ill-equipped to address.

The complexity of humanoid robot-related accidents exceeds that of previous technological innovations in three key characteristics: humanoid simulation, artificial intelligence, and hybrid

[1]*See* JOHN FABIAN WITT, THE ACCIDENTAL REPUBLIC: CRIPPLED WORKINGMEN, DESTITUTE WIDOWS, AND THE REMAKING OF AMERICAN LAW 1–3 (2006).

© The Author(s), 2026. Published by Cambridge University Press on behalf of German Law Journal e.V. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

control. As a result, assigning liability for accidents that occur due to robot malfunctions or erratic behavior becomes a far more complicated task. In cases of humanoid robot accidents, questions arise not only about who is liable—the manufacturer, the programmer, the owner, or the user—but also about how to assess fault when humanoid robots' actions are controlled by both humans and algorithms. As humanoid robots become further integrated into society, China's current accident law system is not yet fully equipped to address the unique risks posed by these technologies. In light of this, a new liability mechanism is needed to fill these gaps.

This Article critically examines these concerns by analyzing the current gaps in Chinese accident law and proposing a new liability mechanism tailored to the challenges posed by humanoid robot accidents.[2] By drawing on lessons from previous technological revolutions, such as those brought about by automobiles, this Article will explore how liability mechanisms can be adapted and restructured to ensure that victims of humanoid robot accidents receive appropriate compensation, while also encouraging innovation in the field of humanoid robotics. In doing so, it will offer a pathway for creating a forward-looking legal regime that not only responds to current challenges but is also flexible enough to accommodate future advancements in this rapidly evolving field.

This Article proceeds by first examining the technological characteristics of humanoid robots that pose unique challenges to existing legal frameworks, specifically focusing on human-body simulation, artificial intelligence, and hybrid control. It then analyzes the limitations of China's current legal system in addressing accidents involving humanoid robots, highlighting the inadequacies of traditional fault standards, the difficulties in identifying algorithmic malfunction, and the shortcomings of current remedies. Following this, the Article presents a typological analysis of humanoid robot accidents, categorizing them based on the source of fault—human or algorithm—to clarify the applicability of existing legal principles. Finally, the Article proposes a reconstructed liability mechanism tailored to the unique challenges of humanoid robot accidents, advocating for new approaches to evidence gathering, algorithmic malfunction regulation, liability allocation, and victim remedies. It concludes by envisioning a path towards an "accidental utopia."

B. Technological Characteristics of Humanoid Robots and Their Influence on Liability

At the 2024 World Robot Conference in Beijing, twenty-seven humanoid robots were showcased.[3] These robots are not only enhancing their ability to perform practical tasks but are also developing the capability to forge emotional connections with humans.[4] China has solidified its role as a global leader in robotics, holding more than 190,000 active patents related to robots as of July 2024—accounting for approximately two-thirds of the world's total.[5] In the past two years, China's Ministry of Industry and Information Technology has issued several executive orders and guidelines to promote the deployment of humanoid robots in sectors such as intelligent manufacturing, domestic services, and hazardous environment operations.[6] As a result, humanoid

[2]*See* Lawrence Lessig, Code: Version 2.0 (2006) (explaining how the law may pose particular functional challenges for the development and deployment of humanoid robots).

[3]*Humanoid Robots Steal the Show at WRC 2024 in Beijing,* cgtn (Aug. 22, 2024), https://news.cgtn.com/news/2024-08-22/Humanoid-robots-steal-the-show-at-WRC-2024-in-Beijing-1wgKidrS7Cg/p.html.

[4]*China's Humanoid Robots Advance with Enhanced Emotional Interaction*, Xinhua News Agency (Aug. 26, 2024), https://english.news.cn/20240826/411a4cec42634f77bd0bf41d36737ddc/c.html.

[5]*See* Sonoda & Kobayashi Intellectual Property Law, *Ministry of Industry and Information Technology: More than 190,000 Effective Patents for Robots in China, Accounting for About Two-Thirds of the World*, Lexology (Sept. 30, 2024), https://www.lexology.com/library/detail.aspx?g=0b9c7206-6b4e-4e33-b2ec-ac887a18568f.

[6]*See, e.g.*, *Guiding Opinions on the Innovation and Development of Humanoid Robots*, Akin Gump Strauss Hauer & Feld LLP (Oct. 20, 2023), https://www.akingump.com/en/insights/ai-law-and-regulation-tracker/guiding-opinions-on-the-innovation-and-development-of-humanoid-robots; *Implementation Opinions on Promoting the Innovative Development of Future Industries*, Akin Gump Strauss Hauer & Feld LLP (Jan. 18, 2024), https://www.akingump.com/en/insights/ai-law-and-regulation-tracker/implementation-opinions-on-promoting-innovation-and-development-of-future-industries.

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

robots, featuring lifelike silicone skin and human-like movements, are becoming increasingly integrated into various public and private sectors in China. Humanoid robots even took to the stage during the 2025 Spring Festival Gala, China's most-watched television show.[7]

As humanoid robots become more prevalent across various sectors, their integration raises new concerns about the potential for accidents,[8] necessitating a careful examination of three critical technological characteristics that must be addressed in any legal framework, namely, human-body simulation, artificial intelligence, and hybrid control.

I. Human-body Simulation

Human-body simulation is the defining hardware feature differentiating humanoid robots from other robots. Humanoid robots are engineered with mechanical structures that closely replicate human sensory organs and limbs, facilitating interactions and operations within human-centric environments, particularly in interacting with human tools. Their stature, often approximating the average human height, further enhances their integration into human society.

Recent advancements in humanoid robotics, exemplified by models such as Engineered Arts' Ameca, Boston Dynamics' Atlas, Tesla's Optimus, Figure AI's Figure 01, and EXRobots' Ex Robots, highlight the rapid evolution of these technologies. Ameca, for instance, demonstrates sophisticated capabilities in facial expression and vocalization, while Atlas exhibits exceptional agility and adaptability to challenging terrains. Figure 01 showcases advanced cognitive functions, including visual perception, planning, and verbal reasoning. Some of these humanoid robots provide even better performance than humans. For example, using machine vision, some medical robots can match photos collected by sensors with those in a pathology database on high-resolution, high-precision electronic grids that are completely imperceptible to the human eye, spotting differences that even experienced specialists might miss.[9]

The rapid advancement of human-body simulation presents a set of legal and ethical challenges. The capacity of these machines to mimic human appearance and behavior enhances their suitability for service roles within human-centric environments, which have been shaped by centuries of human use. First of all, humanoid robots, by virtue of their humanoid form, can seamlessly integrate into these environments, particularly when equipped with bipedal locomotion that allows them to navigate spaces designed for human bodies. Secondly, the vast majority of tools and infrastructure are designed for human use. Humanoid robots can directly utilize these tools, from simple handheld devices to complex medical equipment, enhancing their utility in a variety of tasks. Last but not least, the anthropomorphic design of humanoid robots fosters greater social acceptance and facilitates more natural human-robot interaction. Through facial expressions, gestures, and language, these robots can better understand and respond to human needs, creating a more intuitive and empathetic user experience.[10]

[7]*See* Cheryl Arcibal, *Dancing Kings: Unitree Humanoid Robots, Backed by Alibaba Tech, Delight Spring Gala Show*, South China Morning Post (Jan. 30, 2025), https://www.scmp.com/tech/tech-trends/article/3296720/dancing-kings-unitree-humanoid-robots-backed-alibaba-tech-delight-spring-gala-show.

[8]*See generally* Jack Balkin, *The Three Laws of Robotics in the Age of Big Data*, 78 Ohio St. L.J. 1217 (2017).

[9]Ian Steadman, *IBM's Watson Is Better at Diagnosing Cancer Than Human Doctors*, Wired.com (Feb. 11, 2013), http://www.wired.co.uk/article/ibm-watson-medical-doctor (explaining that, driven by banking and law enforcement, machine vision has achieved superhuman pattern reognition capabilities by rapidly processing and comparing amssive datasets of images). *See also* European Parliament & Council Directive 2024/1689 of June 13, 2024, Laying Down Harmonised Rules on Artificial Intelligence (EC) (explaining how the use of machine vision for emotion recognition is attracting regulatory scrutiny).

[10]*See generally* Christoph Bartneck, Tony Belpaeme, Friederike Eyssel, Takayuki Kanda, Merel Keijsers & Selma Šabanović, Human-Robot Interaction: An Introduction (2nd ed. 2024); Li Zhang, Kamlesh Mistry, Ming Jiang, Siew Chin Neoh & Mohammed Alamgir Hossain, *Adaptive Facial Point Detection and Emotion Recognition for a Humanoid Robot*, 140 Comput. Vision & Image Understanding 93, 95 (2015).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

Even within the relatively controlled environment of industrial settings, robot accidents are not unusual.[11] As humanoid robots become more prevalent in our daily lives, extending into domains such as healthcare, hospitality, and domestic settings, the legal implications of human-robot interactions are poised to expand significantly. These interactions may give rise to a new range of legal issues at the time of accidents. More critically, the nature of these accidents transforms. The replication of human-like motion, sensory perception, and even emotional interaction not only increases the probability of physical harm, as evidenced by existing industrial robot accidents, but also introduces a new dimension of psychological harm. The heightened interaction between humans and humanoid robots can trigger a wide range of emotions, including happiness, sadness, anger, surprise, disgust, embarrassment, excitement, shame, contempt, satisfaction, and amusement,[12] in stark contrast to the more limited interactions typical between humans and traditional robots. Studies demonstrate that individuals can form genuine attachments to these humanoid robots, experiencing feelings of companionship, empathy, and even love.[13] This emotional investment transforms the human perception of robots, from a mere tool to something akin to a social actor, a transformation intensified by the robot's apparent autonomy driven by machine learning algorithms. In other words, human-body simulation, capable of eliciting a wide spectrum of human emotions, may create novel accident scenarios.

As a result, the liability mechanism governing humanoid robots must proactively address not only the increased frequency of accidents but also the qualitatively different and potentially more severe emotional consequences stemming from their human-like design and behavior.

II. Artificial Intelligence[14]

Human-shaped machines may resemble us physically through body simulation, but true humanoid robots require artificial intelligence—the ability to think like humans. Without artificial intelligence, a machine is just a machine, even if it's in a human shape, it can't be called a "robot." It's precisely because it simulates human intelligence that we label it a "robot." Powered by technologies like large language models, machine learning, and data analysis, some of today's humanoid robots exhibit intelligence comparable to, or even exceeding, our own.

[11]*See, e.g.*, Kim Tong-Hyung, *An Industrial Robot Crushed a Worker to Death at a Vegetable Packing Plant in South Korea*, Associated Press (Nov. 9, 2023), https://apnews.com/article/south-korea-robot-worker-death-976dc34487051c762362f17166c17fd7.

[12]*See* John Frank Weaver, *Should AI Makers Be Legally Responsible for Emotionally Manipulating Customers?*, Slate (Jan. 20, 2014), http://www.slate.com/blogs/future_tense/2014/01/20/should_ai_makers_be_legally_responsible_for_emotionally_manipulating_customers.html.

[13]*See* Mustafa Suphi Erden, *Emotional Postures for the Humanoid-Robot Nao*, 5 Int'l J. Soc. Robotics 441, 441 (2013); *see generally* S.M. Mizanoor Rahman & Yue Wang, Dynamic Affection-Based Motion Control of a Humanoid Robot to Collaborate With Human in Flexible Assembly in Manufacturing (2016), https://doi.org/10.1115/DSCC2015-9841; Aryel Beck, Lola Cañamero, Antoine Hiolle & Luisa Damiano, *Interpretation of Emotional Body Language Displayed by a Humanoid Robot: A Case Study with Children*, 5 Int'l J. Soc. Robotics 325, 332 (2013). *See also* Theresa Law, Meia Chita-Tegmark, Nicholas Rabb & Matthias Scheutz, *Examining Attachment to Robots: Benefits, Challenges, and Alternatives*, 11 ACM Transactions on Hum.-Robot Interaction 1 (2022) (surveying HRI literature documenting humans' emotional attachment and companionship-oriented bonds with social robots); Elena Morgante, Carla Susinna, Laura Culicetto, Angelo Quartarone & Viviana Lo Buono, *Is it Possible for People to Develop a Sense of Empathy Toward Humanoid Robots and Establish Meaningful Relationships With Them?*, 15 Front. Psychol. 1 (2024) (systematic review finding that robots with anthropomorphic and emotion-responsive features can evoke empathy and facilitate emotional connection).

[14]The concept of artificial intelligence in this Article does not distinguish software and hardware. This is a growing trend in leading AI regulation policies. *See, e.g.*, *AI Strategy 2019: AI for Everyone: People, Industries, Regions and Governments*, Japnese Government Integrated Innovation Strategy Promotion Council Decision 2-3 (T.a.I. Council for Science trans., 2019), https://www8.cao.go.jp/cstp/ai/aistratagy2019en.pdf; *Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector*, Monetary Authority of Sigapore 5 (Nov. 12, 2018), https://www.mas.gov.sg/-/media/mas/news-and-publications/monographs-and-information-papers/feat-principles-updated-7-feb-19.pdf.

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

From a computer science perspective, machine learning is a central enabling technology for the forms of AI most relevant to contemporary humanoid robots. To understand why, it is instructive to examine the limitations of the counterpart of machine learning algorithms—algorithms without machine learning capabilities, referred to as "traditional algorithms." While traditional algorithms have long been the bedrock of computer programming, providing precise and predictable outcomes, they suffer from a fundamental inflexibility. Designed to operate within strictly defined parameters, traditional algorithms struggle to adapt to novel or unexpected circumstances. This rigidity is a direct consequence of their deterministic nature, where every output is strictly determined by the input and the predefined rules of the algorithm. In contrast, machine learning algorithms, through their ability to learn from data, can adapt to changing environments and perform tasks that would be prohibitively complex for traditional programming approaches.[15]

Take the navigation system of a humanoid robot as an example. Traditional algorithmic approaches to robotic navigation, while effective in highly controlled environments, face significant limitations when confronted with the complexities of real-world scenarios.[16] Consider the task of guiding a humanoid robot through a bustling city intersection. While a conventional algorithm might successfully navigate a factory floor with predictable obstacles, it would struggle to account for the dynamic and unpredictable nature of urban traffic, including pedestrians, cyclists, and vehicles. This is because traditional algorithms rely on a predefined set of rules and conditions, making them ill-equipped to handle unforeseen circumstances. Such limitations can be attributed to what is known as the "knowledge acquisition bottleneck,"[17] where the algorithm's ability to adapt to new situations is constrained by the programmer's ability to anticipate and code for every possible eventuality. This inherent inflexibility has hindered the widespread deployment of robots in complex, real-world environments.

Machine learning algorithm has fundamentally changed how humanoid robots function.[18] Unlike traditional algorithm, machine learning algorithm allows humanoid robots to learn from experience, adapt to changing circumstances, and make independent decisions.[19] In a nutshell, machine learning algorithm is both the program itself and the programmer who constantly modifies itself. Each new run of the machine learning algorithm serves as a steppingstone for the next run, improving through the continuous accumulation of operational data. Most importantly, machine learning algorithms will bring uncertainty to the behavior of humanoid robots. As Ryan Calo has said, one of the important characteristics of machine learning algorithms is that they can produce innovative and unpredictable outputs, which he calls "emergence."[20] It is precisely through "emergence" that machine learning algorithms can help humanoid robots cope with dynamic scenarios that traditional algorithms cannot handle. Of course, "emergence" is not

[15]Council Directive 2024/1689, at 12
("This capability to infer refers to the process of obtaining the outputs, such as predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to a capability of AI systems to derive models or algorithms, or both, from inputs or data . . . . The capacity of an AI system to infer transcends basic data processing by enabling learning, reasoning or modelling.")

(according to EU AI Act, the key characteristic of AI systems is "their capability to infer").

[16]Traditional algorithms are only suitable for highly controllable environments. Only in this way can programmers predict the limited number of scenarios that the robot may encounter, design corresponding "if-then-else" code rules, and execute navigation control accordingly. *See* Ajay Agrawal, Joshua Gans & Avi Goldfarb, *The Simple Economics of Machine Intelligence*, Harv. Bus. Rev., Nov. 17, 2016, https://hbr.org/2016/11/the-simple-economics-of-machine-intelligence (explaining that).

[17]Pedro Domingos, The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World 89–90 (2015).

[18]While humanoid robots still rely on traditional algorithms like any machine, it's their use of machine learning algorithms that sets them apart, especially when considering legal liability.

[19]*See* David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn about Machine Learning*, 51 U.C. Davis L. Rev. 653, 672–77 (2017).

[20]*See* Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 Calif. L. Rev. 513, 538–45 (2015).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

without its threshold and costs. The black-box nature of deep learning models, particularly in unsupervised and reinforcement learning settings, makes it difficult, sometimes impossible, to predict and explain the behavior of humanoid robots.[21] This lack of transparency poses profound legal and ethical implications, as it becomes increasingly challenging to assign liability for the actions of autonomous machines.

The integration of artificial intelligence into humanoid robots introduces a new layer of complexity to traditional liability mechanisms. This complexity arises from two primary sources: the multifaceted nature of AI development and the inherent unpredictability of machine learning algorithms. On the one hand, unlike traditional machines, humanoid robots often involve multiple developers and suppliers, each contributing to the hardware, software, and algorithms that underpin the robot's functionality.[22] This distributed development process can make it challenging to pinpoint a single entity responsible for a malfunction or accident. Beyond establishing culpability, these varied scenarios highlight the broad spectrum of applicable legal principles, such as negligence, product liability, and contractual obligations, including warranties, consumer responsibilities, and liability restrictions.[23] On the other hand, machine learning algorithms can exhibit emergent behaviors, meaning that they can produce outcomes that were not explicitly programmed. This unpredictability makes it difficult to anticipate and prevent accidents, and it can also complicate the task of assigning liability after an incident. These factors combine to create significant uncertainty regarding the allocation of liability in cases involving humanoid robots. Traditional liability mechanism, which typically relies on relatively clear causal connections between a defendant's conduct and a resulting harm, is inadequate for addressing the complex and often unpredictable nature of humanoid robot accidents.[24]

III. Hybrid Control

Despite the advancements in artificial intelligence, at this stage, humanoid robots remain subject to human control. Even in the most sophisticated of these machines, human intervention remains a critical component, from the initial programming and activation to the ultimate ability to deactivate the system. This human-in-the-loop paradigm is evident in the three-step process typically involved in human-robot interaction: instruction, reception, and execution.[25] While machine learning algorithms grant robots a degree of autonomy in carrying out tasks, humans retain the overarching authority to initiate actions, modify behaviors, and ultimately terminate operations.[26] This complex interplay between human operators and autonomous algorithms in humanoid robots is another defining characteristic of humanoid robotics.

From a legal liability perspective, the determination of causality within human-algorithm hybrid control systems presents an exponential increase in complexity.[27] Questions arise regarding whether a harmful event was caused by algorithmic failure, human error, or external

[21] *See id.* at 538–39. *See generally* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2016) (discussing algorithmic transparency).

[22] *See* Jennifer Cobbe & Jatinder Singh, *Artificial Intelligence as a Service: Legal Responsibilities, Liabilities, and Policy Challenges*, 42 COMPUT. L. & SEC. REV. 1, 3–7 (2021).

[23] Charlotte A. Tschider, *Humans Outside the Loop*, 26 YALE J. L. & TECH. 324, 356 (2024).

[24] *See generally* Jan De Bruyne & Wouter Ooms, *Tort Liability and Artificial Intelligence*, *in* THE CAMBRIDGE HANDBOOK OF THE LAW, ETHICS AND POLICY OF ARTIFICIAL INTELLIGENCE 158 (2025).

[25] J. E. Dominguez-Vidal, Nicolas Rodriguez, Rene Alquezar & Alberto Sanfeliu, *Perception-Intention-Action Cycle in Human-Robot Collaborative Tasks*, ARXIV (June 1, 2022), https://arXiv.org/abs/2206.00304.

[26] *See generally* Hamed Bozorgi & Trung Dung Ngo, *Beyond Shared Autonomy: Joint Perception and Action for Human-In-The-Loop Mobile Robot Navigation Systems*, 109 J. INTELLIGENT & ROBOTIC SYS. 19 (2023).

[27] To clarify hybrid control in human-algorithm interaction, some research classifies medical robot accident liability into three dimensions: human-aided incidents, where robots are used as tools—programming errors—technical malfunctions—and autonomous decision-making, where robots operate independently. *See generally* Xiaoli Shentu, *A Review on Legal Issues of Medical Robots*, NAT'L LIBR. MED., May 24, 2024, https://pubmed.ncbi.nlm.nih.gov/38788002/. However, such

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

factors. The intricate relationship between human input, algorithmic decision-making, and the resulting outcome makes it difficult to establish a clear causal chain. Furthermore, the emergent properties of machine learning algorithms, which can produce unexpected results, further complicate the task of assigning liability.

The unpredictability inherent in machine learning systems that are being built in an increasing number of humanoid robots challenges traditional notions of foreseeability and culpability. As a result, the legal system must grapple with novel questions, such as: who is liable when a machine learning algorithm makes a decision that results in harm? How can we determine the extent to which a human operator should be held responsible for the actions of a humanoid robot? And how can we ensure that victims of harm have adequate remedies when the causes of accidents are so difficult to understand?

To sum up, the ongoing advancement of humanoid robotics, characterized by the convergence of physical embodiments and increasingly sophisticated artificial intelligence, reflects a trajectory reminiscent of the Tyrell Corporation's famous motto in *Blade Runner*: "More human than human." However, despite these advancements, the current state of humanoid robotics still represents a transitional phase. While exhibiting increasing autonomy, these machines remain subject to a complex interplay of human control and algorithmic decision-making, thereby complicating the attribution of liability for their actions in accidents.

C. Limitations of the Traditional Legal System in China

In the digital age, humans frequently encounter errors when using algorithms. A common example is a writer losing substantial work due to a software crash, despite the existence of an auto-save function. Even if the software's auto-save feature is flawed, lawsuits against the software company are rare. This situation highlights the limitations of tort law in addressing algorithmic failures.[28] Typically, liability serves as a secondary recourse, with contractual terms, such as those found in software user agreements, governing the primary legal relationship between the user and the service provider. In practice, courts are hesitant to recognize tort claims based solely on algorithmic errors, often requiring additional elements like harm to third parties or personal injury.[29] However, the increasing prevalence of humanoid robots presents a new challenge. As robots become more integrated into human society and take on tasks traditionally performed by humans,[30] the potential for harm resulting from algorithmic failures grows significantly. When such failures cause personal injury or property damage, courts are more likely to entertain tort claims, as contract law is insufficient to deal with these accidents.

The Chinese liability mechanism, while undergoing modernization, still grapples with the complexities of emerging technologies, particularly with regard to humanoid robots. Traditional legal concepts within Chinese civil law, such as strict liability and fault-based liability, may prove insufficient to address the nuanced issues presented by accidents involving these robots. Specifically, the unique characteristics of humanoid robots generate at least three distinct challenges for the existing Chinese liability mechanism.

I. Standard of Fault

The standard of "fault" as it applies to humanoid robots is fundamentally different from that of traditional machines. In accidents involving traditional machines, where product liability was not

classifications do not adequately address the complexities of hybrid control, where identifying the responsible agent at the moment of an accident becomes problematic.

[28]*See* Emily Kuwahara, Note, *Torts v. Contracts: Can Microsoft Be Held Liable to Home Consumers for Its Security Flaws?*, 80 S. CAL. L. REV. 997, 998 (2007); DAN B. DOBBS, THE LAW OF TORTS § 114, 269 (2001).

[29]*See* Michael D. Scott, *Tort Liability for Vendors of Insecure Software: Has the Time Finally Come?*, 67 MD. L. REV. 425, 430 (2008).

[30]*See infra* Section D.2.1.

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

applied, human error is often the primary focus, with legal doctrines like negligence and emergency defense being designed to address situations where a human operator has made a mistake.[31]

Unlike traditional machines, humanoid robots exhibit a degree of autonomy enabled by machine learning algorithms. This autonomy, coupled with the intricate nature of human-algorithm hybrid control, makes it challenging to pinpoint the precise cause of an accident and to establish the standard of fault. For instance, consider a scenario where a humanoid robot, while performing household chores, accidentally damages a valuable object or injures a child. In such cases, traditional liability analysis, which might focus on the operator's level of care, may be insufficient. The robot's actions are influenced by a multitude of factors, including the operator's instructions,[32] the robot's internal algorithms, and real-time data processing. As a result, if a humanoid robot causes harm, it may be difficult to determine whether the fault lies with the human operator, a flaw in the machine learning algorithm, or a combination of both. Furthermore, the appropriate boundaries of negligence in hybrid control circumstances remain uncertain.[33] Traditional Chinese accident law, primarily based on conventional product liability principles, fails to adequately address the complexities of hybrid control circumstances, particularly the shared fault and potential joint liability arising from human-algorithm hybrid control. Lastly, the intricacies of human-robot interaction, including the potential for miscommunication or misunderstanding, further complicate the issue. Human-robot interactions are not merely input-output exchanges, but rather involve a dynamic interplay of human intent, algorithmic interpretation, and real-time adaptation, making it exceptionally challenging to isolate a single, definitive causal factor. To sum up, unlike traditional machines, where the operator is typically in direct control of the machine's actions, humanoid robots can exhibit a degree of autonomy that makes it difficult to establish the standard of fault.

II. Algorithmic Malfunction

Even if the liability mechanism works on a theoretical and textual level, the practical effectiveness of the liability mechanism to ensure compensation of the victim rests upon how readily victims can establish liability in adjudication. To that end, the evidentiary mechanism for identifying algorithmic malfunction becomes another major challenge to the existing liability mechanism.

As discussed above, the increasing delegation of control to artificial intelligence systems has made the behavior of humanoid robots more unpredictable. This unpredictability is exacerbated in dynamic environments where robots must continuously adapt to changing conditions, potentially leading to unforeseen behaviors that neither the operator nor the algorithm developers could have reasonably anticipated. Indeed, in many instances, neither the human user nor the programmer who developed the algorithm can fully comprehend the automated decision-making process.[34] This is not unique to humanoid robots; it is a common characteristic of machine learning algorithms in general.

Consider the case of AlphaGo, the renowned machine learning algorithm that defeated top human Go players in 2017. Even the engineers who created AlphaGo could not fully explain the intricacies of its moves—otherwise the engineers themselves would be the world champions. This highlights the inherent challenge of interpreting and understanding the

[31]Jídù wéixiǎn huódòng de zérèn (极度危险活动的责任) [Liability for Ultra-hazardous Activities] (Civil Code of the People's Republic of China, adopted by the Third Session Nat'l People's Cong., effective on January 1, 2021) ch. 8, art. 1240, 223 (China).

[32]*See, e.g.*, Andrea Bertolini, Pol'y Dep't for Citizens' Rts. & Const. Affs., Artificial Intelligence and Civil Liability 52 (2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf (noting that there may be front-operators and back-operators).

[33]*See generally* Andrew D. Selbst, *Negligence and AI's Human Users*, 100 B.U. L. Rev. 1315 (2020).

[34]*See* Domingos, *supra* note 17, at 258–59.

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

decision-making processes of complex algorithms. If even in the highly controlled environment of a 19×19-grid Go board, engineers struggle to explain the algorithm's choices, the difficulty is magnified exponentially in the complex and unpredictable scenarios encountered by humanoid robots in real-world service settings. The Boeing 737 MAX crashes of 2018 and 2019 serve as a stark reminder of this issue.[35] The aircraft's Maneuvering Characteristics Augmentation System (MCAS), a flight control system, was implicated in both accidents.[36] Despite extensive investigations, Boeing engineers and the official investigation teams took a long time to pinpoint and agree to the exact nature of the algorithmic malfunction that led to the crashes.[37]

Humanoid robots, like modern aircraft, operate in complex environments where various factors, such as user input, environmental changes, and external influences, can impact algorithmic decision-making. Consequently, pinpointing algorithmic malfunction presents a significant challenge, hindering the establishment of a factual basis for legal evidence. This difficulty is compounded by inherent informational asymmetry: mechanical information is typically controlled by the robot's manufacturer, while those seeking compensation for damages often face a substantial disadvantage in accessing and understanding the product's design and operation. This disparity significantly impedes the evidence-gathering process necessary to identify algorithmic malfunction.

III. Remedies for Humanoid Robots' Damages

The traditional tort system faces significant challenges in providing effective redress for damages caused by malfunctioning humanoid robots. Machine learning algorithms, which empower these robots to perform tasks autonomously, introduce the possibility of robots initiating harmful actions. While traditional accident law offers avenues for financial compensation by assigning legal responsibility to the robot's developer or user,[38] it falls short when attempting to rectify the harmful behavior itself. As Mark Lemley and Bryan Casey have observed, traditional legal remedies like punitive damages or injunctions are designed to deter human behavior; however, traditional deterrents such as imprisonment or financial penalties are inapplicable to humanoid robots.[39]

Furthermore, while some scholars have begun exploring legal remedies and related rules for robots,[40] comparatively little attention has been paid to the psychological remedies. Unlike humans, humanoid robots lack moral agency, emotions, and the capacity to understand punishment. However, unlike traditional machines, humanoid robots feature human-body simulation and can elicit emotional responses in human victims. Consequently, victims of harm might derive some psychological relief from physically interacting with or damaging the offending robot. These considerations present a novel challenge, further complicated by the rapid advancement of humanoid robot technology and the need to anticipate future scenarios.

[35]Federal Aviation Administration, Summary of the FAA's Review of the Boeing 737 MAX 1–2 (2020), https://www.faa.gov/sites/faa.gov/files/2022-08/737_RTS_Summary.pdf.

[36]*Id.*

[37]*See generally* Nat'l Transp. Safety Bd., Response to Final Aircraft Accident Investigation Report: Ethiopian Airlines Flight 302 (2019), https://www.ntsb.gov/investigations/Documents/Response%20to%20EAIB%20final%20report.pdf.

[38]*See* S nhài péicháng (损害赔偿) [Damages] (Civil Code of the People's Republic of China, effective on January 1, 2021) ch. 2, arts. 1179–82 (China).

[39]*See* Mark A. Lemley & Bryan Casey, *Remedies for Robots*, 86 U. Chi. L. Rev. 1311, 1361–67 (2019).

[40]*See generally id.*; *see also* Maayan Perel, *Digital Remedies*, 35 Berkeley Tech. L.J. 1, 47 (2020).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

D. Typological Analysis of Humanoid Robot Accidents

In Chinese law, accident liability determination continues to rely substantially on fault-based reasoning.[41] A common example is road traffic accidents.[42] Article 76 of the Road Traffic Safety Law establishes the following liability mechanism: in accidents between vehicles, a fault-based liability mechanism is implemented; and if both parties are at fault, they bear responsibility proportionally. In accidents between vehicles and non-motor vehicles or pedestrians, the principle of presumed fault and no-fault liability is implemented.

However, as previously mentioned, the presence of artificial intelligence and human-machine hybrid control complicates fault determination in humanoid robot accidents. According to the subject of fault of the humanoid robot—human or algorithm—we can divide humanoid robot accidents into the following two categories: human fault and robot fault.

I. Human Fault

Under the current Chinese liability mechanism, human fault encompasses both intentional acts and negligence, including a distinct category of "gross negligence."[43]

1. Intentional Acts

Intentional acts can be further analyzed in two scenarios. The first involves the deliberate use of humanoid robots to cause harm, such as in cases of murder, assault, or property damage. These actions clearly lie outside the scope of accident law. The second scenario encompasses the infliction of *deliberate least-cost harm*, where an actor, confronted with a choice between two undesirable outcomes, intentionally chooses the lesser of two evils. This scenario, often likened to the trolley problem,[44] assumes the individual possesses sufficient time for reasoned decision-making. Traditional accident law addresses such situations through doctrines such as "inevitable necessity,"[45] which is also recognized within Chinese accident law.[46] Crucially, neither of these intentional acts presents novel legal challenges uniquely attributable to the use of humanoid robots.

2. Negligence

Human negligence involving humanoid robot accidents may initially appear analogous to negligence in traditional mechanical accidents, often arising from carelessness or overconfidence.[47] Recognizing the opaque nature of human cognition, accident law employs the reasonable person standard as the touchstone for evaluating negligence, or gross negligence.[48] However, the increasing sophistication of humanoid robots may elevate the standard of care expected of their users, thereby

[41]Yang Lixin, 我国道路交通事故责任归责原则研究 [*Research on Traffic Accident Liability in China*], 10 法学 [L. Sci.] 109, 109–18 (2008).

[42]*See* Han Xuzhi, 自动驾驶事故的侵权责任构造 [*The Liability Structure of Autonomous Driving Accidents*], 上海大学学报(社会科学版) [Shanghai Univ. J.], Mar. 2019 (Soc. Sci. ed.), https://m.qikan.cqvip.com/Article/ArticleDetail?id=00002EGMK9707JP167DO3JP1MDR&from=Article_ArticleDetail (providing a relevant theoretical review).

[43]Gross negligence, a significantly higher degree of carelessness or recklessness than ordinary negligence, can result in more serious legal repercussions. *See* Ye Mingyi, *The Construction of Gross Negligence Theory*, 31 Chinese J.L. 77, 77–90 (2009).

[44]*See* Judith Jarvis Thomson, *Killing, Letting Die, and the Trolley Problem*, 59 Monist 204, 206 (1976).

[45]*See* Stephen G. Gilles, *Inevitable Accident in Classical English Tort Law*, 43 Emory L.J. 575, 577 (1994); Richard W. Wright, *Negligence in the Courts: Introduction and Commentary*, 77 Chi.-Kent L. Rev. 425, 480 (2002).

[46]*See* Fu Qiang, *On Civil Liability of Necessity*, 4 Zhejiang Acad. J. 150, 150–56 (2010); Liu Zuoxiang, *Necessity: Institutional Design for Rights Conflict Resolution and the Related Criminal and Civil Cases*, 1 Hebei L. Sci. 12, 12–21 (2014); Song Zongyu & He Wenhao, *The Practice Forms and Normative Interpretation of Good Samaritan Rule in Civil Code*, 6 Qilu J. 75, 75–87 (2023).

[47]*See* Richard W. Wright, *The Standards of Care in Negligence Law*, *in* Philosophical Foundations of Tort Law 249 (David G. Owen ed., 1995).

[48]Consider the "4D" framework for negligence for human drivers: Drunk, Drugged, Drowsy, and Distracted. These states often impair an individual's judgment and lead to actions that deviate from the reasonable person standard.

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

expanding the scope of what constitutes negligence. Furthermore, delegating tasks to a humanoid robot does not automatically diminish the user's responsibility. For instance, a driver excessively reliant on autonomous driving features may be more inclined to engage in distracting activities, such as reading, sleeping, or using a mobile phone, rather than diligently monitoring the vehicle's operation. Consequently, humanoid robots often incorporate enhanced warning systems to mitigate the risks associated with user inattention or misuse. Nevertheless, these considerations impact the apportionment of liability at the margins; they do not fundamentally alter the nature of human negligence in this context. Therefore, instances of human negligence in humanoid robot accidents remain amenable to analysis within the established framework of traditional accident law.

II. Robot Fault

Robot fault, distinct from human fault which stems from human fallibilities, manifests in three primary forms, each rooted in the robot's algorithmic programming and operational design.

1. Traditional Algorithmic Defects

This category encompasses traditional algorithmic flaws similar to those found in conventional robots, which are not equipped with machine learning algorithms—examples include warning system failures, unresponsive controls, user privacy breaches, and cybersecurity vulnerabilities. These traditional algorithmic defects in humanoid robots can be treated analogously to hardware defects of conventional robots. When a humanoid robot's hardware design is flawed or fails to meet industry technical standards, principles of product liability may be applicable.[49] Consequently, no substantial modifications to existing accident law are required to address these traditional algorithmic defects.

2. Machine Learning Algorithmic Defects

This category involves defects specific to the machine learning algorithms employed by humanoid robots. These defects can manifest during the data collection process—for example, insufficient image resolution, incorrect object recognition—or during data analysis and processing—miscalculation of outputs, for example. Ultimately, these defects can lead to erroneous automated decision-making, inappropriate algorithmic intervention, and, consequently, accidents. Furthermore, humanoid robots may present systemic harm issues stemming from machine learning algorithms. For example, systemic discrimination arising from machine learning algorithms can result in differentiated safety standards for users based on factors such as income level or gender.[50] These issues fall within the scope of machine learning algorithmic defects, a type of algorithmic malfunction for which traditional accident law struggles to provide adequate solutions. Therefore, addressing these novel challenges may necessitate the development of a new liability mechanism specifically tailored to the technical features of machine learning algorithms.

3. "Intentional" Algorithmic Act

This category involves machine learning algorithms that "intentionally" cause accidents.[51] The objective of such robot fault is to minimize harm or to achieve deliberate least-cost harm, but the

[49]*See generally* Product Liability (Civil Code of the People's Republic of China, effective on January 1, 2021) STANDING COMM. NAT'L PEOPLE'S CONG. GAZ., ch. 4, sec. 7 (China).

[50]*See* Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 1, 10-18 (2014); Andrew Tutt, *An FDA for Algorithms*, 69 ADMIN. L. REV. 83, 105 (2017); Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 690 (2016).

[51]The "intent" of machine learning algorithms is a functional concept, not a substantive one. Like the concept of a legal person or a property right, it is created solely to serve the allocation of accident liability. It does not mean that the algorithm

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

"intent" demonstrated by the algorithm is more formally rational than human intent. Technically, algorithms differ from humans in their information processing speed. Even in a fraction of a second before an accident, an algorithm can perform rational calculations and make automated decisions.[52] In other words, it is a "decision," rather than a "reaction." Therefore, seemingly accidental events may, in fact, be the result of a rational choice made by the algorithm,[53] exhibiting a level of formal logic exceeding that of a human. Although the rational choices of machine learning algorithms are typically based on correlation, rather than causation,[54] in emergency situations involving the control of humanoid robots, the automated decisions made by the algorithm approach human intent. This may be understood as a functionally intentional form of automated decision-making in emergency situations, as distinguished from "intentional" automated decision-making resulting from programmer intervention or traditional algorithmic programming.

A related issue, akin to the trolley problem, is particularly salient: to avoid harm to the user, the algorithm "intentionally" harms a third party as a sacrifice.[55] In an imminent accident scenario, the humanoid robot, based on the information that it learned, faces a choice: protect the user or third parties—other persons or property? Autonomous driving manufacturers, driven by commercial interests, naturally tend to prioritize the protection of users in their algorithm design. From a business model perspective, users are direct customers, while third parties are not.[56] Moreover, considering the information cost, the sensors within humanoid robots struggle to provide reliable judgments concerning dynamic external scenarios—such as the quantification, resistance, and evasive capabilities of third parties—in contrast to the relatively predictable nature of human user responses. As a result, the occurrence of a humanoid robot algorithm "intentionally" causing harm to a third party in unavoidable circumstances becomes more probable. The challenge then arises: how can a third party, who sustains personal injury or property damage from such an algorithm's "intentional" harm, seek legal redress? This scenario exposes a gap in the conventional accident liability mechanism.

E. Reinventing a Liability Mechanism for Humanoid Robot Accidents

The preceding classification of humanoid robot accidents distinguishes between two sets of interacting agents: natural persons and algorithms, and traditional algorithms and machine

truly possesses the same subjective intent as humans. At least in the current stage of artificial intelligence development, robots with autonomous consciousness like those in science fiction works such as *Westworld* and *Blade Runner* do not yet exist, nor do they need to be included in the scope of accident law that we discussed in this Article.

[52]For example, assuming a humanoid robot must choose between causing death and causing injury, a machine learning algorithm, through continuous learning, makes the following cost-benefit judgment: compared to causing injury, the average cost required for causing death is lower—this cost is only for the manufacturer, not for the whole society, and thus continuously strengthens the automated decision-making of causing death.

[53]*See generally* Ryan Abbott, The Reasonable Robot: Artificial Intelligence and the Law (2020).

[54]*See generally* Viktor Mayer-Schönberger & Kenneth Cukier, Big Data: A Revolution That Will Transform How We Live, Work, and Think (2014).

[55]*See, e.g.*, Matt Simon, *To Make Us All Safer, Robocars Will Sometimes Have to Kill*, Wired.com (Mar. 13, 2017), https://www.wired.com/2017/03/make-us-safer-robocars-will-sometimes-kill/, (showing an early discussion on the issue of algorithms intentionally harming a third party to avoid harm the user in autonomous driving technology); *see generally* Ryan Calo, *Is the Law Ready for Driverless Cars?*, 61 Commc'ns ACM 34, 35 (2018); Anna Jobin, Marcello Ienca, & Effy Vayena, *The Global Landscape of AI Ethics Guidelines*, 1 Nature Mach. Intel. 389 (2019); Chen Jinghui, *Autonomous Driving and Passenger Priority*, J. E. China Univ. Pol. Sci. & L., No. 6, 2020; Zheng Yushuang, *Algorithmic Justice and Legal Responsibility System for Autonomous Driving,* Legal System and Social Development, No. 4, 2022.

[56]*See* Michael Taylor, *Self-Diving Mercedes-Benzes Will Prioritize Occupant Safety over Pedestrians*, Car & Driver (Oct. 7, 2016), https://www.caranddriver.com/news/a15344706/self-driving-mercedes-will-prioritize-occupant-safety-over-pedestrians/ (reporting that Christoph von Hugo, head of Mercedes-Benz's autonomous driving safety department, admitted in an interview that if there is a choice between the two, Mercedes-Benz cars will prioritize protecting the driver in the car rather than people outside the car).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

Source of Fault	Types of Fault	Legal System
Human Fault	Intentional Act	Traditional Accident Law
	Negligence	Traditional Accident Law
Robot Fault	Traditional Algorithmic Defects	Traditional Product Liability Law
	Machine Learning Algorithmic Defects	?
	Intentional Algorithmic Act	?

Figure 1. Typology of liability constellations

learning algorithms. This typological approach serves not as an end in itself, but as a means to elucidate the layered legal issues arising from humanoid robot accidents, ultimately determining which categories are amenable to traditional accident liability mechanism and which suggest the need for a new mechanism.

Through this typological analysis (Fig. 1), the feasibility and limitations of traditional accident liability mechanisms in the context of humanoid robot accidents become apparent. Traditional liability mechanisms remain applicable only in situations where human fault and standard algorithmic defects are present, and even then, contingent upon the clear demarcation of responsibility between human and robot.

Traditional accident liability mechanism proves inadequate for addressing the complex scenarios involving the commingling of human and machine fault, machine learning algorithmic defects, and, especially, accidents "intentionally" caused by machine learning algorithms. Addressing these novel challenges posed by humanoid robots is the central focus of constructing a robust and effective liability mechanism for such accidents. The following sections will explore the development of such a system, concentrating on the issues of commingled human-machine fault, machine learning algorithmic defects, accidents "intentionally" caused by these algorithms, and remedies for humanoid robot harms.

I. Establishing Evidentiary Process for Determining the Legal Facts of Hybrid Control

The foregoing classification of humanoid robot accidents, distinguishing between natural persons and algorithms, as well as traditional and machine learning algorithms, reveals the limitations of traditional accident liability mechanisms. While these frameworks adequately address scenarios involving human fault and traditional algorithmic defects, they falter when confronted with the complexities of commingled human-machine fault, machine learning algorithmic defects, and "intentional" robot actions. Without a clear understanding of who or what is at fault—human, algorithm, or a combination thereof—meaningful liability cannot be assigned. Therefore, establishing a robust evidentiary process is crucial to support the effective operation of a liability mechanism for humanoid robot accidents.

The transition from exclusive human control to human-machine hybrid control, coupled with the inherent uncertainty of machine learning algorithms in the determination of legal facts, poses a significant challenge to traditional fault-based accident liability mechanisms.[57] Distinguishing between user misoperation and algorithmic defects, for example, often proves difficult in practice.

[57]*See* Qīnquán zérèn—yībān guīzé (侵权) [Tort Liability—General Rules] (Civil Code of the People's Republic of China, effective on January 1, 2021) book 7, ch. 1, art. 1168 (China) ("[W]here two or more persons jointly commit a tortious act causing damage to another person, they shall bear joint and several liability.") [hereinafter Tort Liability]. This doctrinal joint liability mechanism, as well as traditional theories of joint liability, is insufficient to address the problem of humanoid robot accidents. They presuppose that the actors involved are all human beings, and do not explore how algorithms, especially machine learning algorithms, can share responsibility with humans.

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

Similar challenges pervade the determination of control in human-machine hybrid control, a critical technical characteristic that cannot be circumvented in the regulation of humanoid robots.

How can we determine whether the dominant control in human-machine hybrid control at the time of the accident resided with the user or the machine learning algorithm? Even in situations where the user appears to be in control, did the machine learning algorithm faithfully execute the user's commands? These questions present significant obstacles to factual determination for legal adjudication. For example, in disputes concerning autonomous driving accidents, the central point of contention frequently revolves around whether the proximate cause of the accident was a command originating from the user or the algorithm.[58]

Therefore, regulatory authorities should implement *ex ante*, *in situ*, and *ex post* measures to establish the legal facts regarding the existence of machine learning algorithmic defects and whether the machine learning algorithm "intentionally" caused the accident. *Ex ante*, regulatory agencies should mandate that manufacturers regularly file information pertaining to the functions of their machine learning algorithms, the processes of algorithmic automated decision-making, and other relevant data.[59] *In situ*, the operation of humanoid robots should be monitored by a dedicated safety monitoring platform to track algorithmic operational information. Manufacturers should be incentivized to appropriately retain operational data and logs, while ensuring the protection of personal information and data security, to facilitate evidence collection in potential accidents. Analogous to established practices in civil aviation and rail transport, humanoid robots require an Event Data Recorder (EDR) system—a "black box"—which, upon triggering specific conditions, meticulously records data including speed, location, time, safety system status, algorithmic control actions, human user control actions, and other pertinent information. *Ex post*, regulatory agencies should initiate third-party algorithmic testing and auditing.[60] This measure is aimed not only at tracing the root cause and revealing design or execution flaws in the machine learning algorithm but also at evaluating the causal link between any such flaws and the accident.[61] Critically, victims seeking compensation for damages often face a significant disadvantage compared to manufacturers regarding access to and understanding of product design and operational information. This informational asymmetry can impede the fair apportionment of risk, especially in cases involving technical complexity.[62] Therefore, facilitating victims' access to evidence for legal proceedings, especially at the *ex post* stage, is essential for legal remedies.

II. Regulating Algorithmic Malfunction Through the "Reasonable Person" Technical Standard

Given the inherent potential for accidents arising from the use of humanoid robots, the objective of a liability mechanism should adopt a more pragmatic approach, shifting from the idealistic goal of complete accident *prevention* to the more realistic aim of accident *mitigation*.[63] This pragmatic philosophy can be traced back to the origins of traffic accident law, notably through the work of Ralph Nader, whose book, *Unsafe at Any Speed*,[64] profoundly shaped the American traffic

[58]*See* Elsie Chen, *Her Parents Were Injured in a Tesla Crash. She Ended Up Having to Pay Tesla Damages*, YAHOO! FINANCE (Feb. 11, 2025), https://finance.yahoo.com/news/her-parents-were-injured-tesla-010111030.html (reporting how one Tesla crash case in China required several years to resolve evidentiary issues concerning the attribution of fault between the autonomous driving system and the driver).

[59]*See* Tutt, *supra* note 50, at 118–23.

[60]*See* Ding Xiaodong, *On the Legal Regulation of Algorithms*, 17 FRONTIERS L. CHINA 88, 88 (2022).

[61]While the "black box" nature of algorithms raises legitimate concerns, recent scholarship suggests that algorithms may, in fact, be more readily monitored and controlled than the human mind. *See* DANIEL KAHNEMAN, OLIVIER SIBONY & CASS R. SUNSTEIN, NOISE: A FLAW IN HUMAN JUDGMENT (2021).

[62]*See* Directive 2024/2853 of Oct. 23, 2024, on Liability for Defective Products and Repealing Council Directive 85/374/EEC, 2024 O.J. (L 2853) 2 (EC).

[63]Bryan H. Choi, *Crashworthy Code*, 94 WASH. L. REV. 39, 86–91 (2019).

[64]*See generally* RALPH NADER, UNSAFE AT ANY SPEED: THE DESIGNED-IN DANGERS OF THE AMERICAN AUTOMOBILE (1965).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

accident liability mechanism and influenced international discourse. Under Nader's influence, the dominant adjudicative principle became that vehicle manufacturers are not obligated to produce flawless vehicles, but are bound by a duty to "exercise reasonable care in the design and manufacture of their product to minimize the personal injuries to which its users are exposed; and to not place its users in an unreasonable risk of personal injury when a collision occurs."[65] The concepts of "reasonable care" and "unreasonable risk of personal injury" effectively integrate the standard of reasonableness into the product design, development, and manufacturing processes. This form of *ex-ante* regulation, framed as a safety standard, is not merely beneficial but essential for products like humanoid robots, which are capable of actively initiating harm. To effectively address humanoid robot accidents, we must consider specific technological realities, weigh the costs of technological implementation, and adopt reasonable *ex-ante* regulatory mechanisms to control algorithmic malfunction, thereby reducing the incidence of such accidents.

Crucially, technical standards inherently incorporate fault-based liability standards from accident law.[66] To assess user negligence, specifically, whether the user exercised reasonable care, the traditional accident liability mechanism employs the "reasonable person" standard, the baseline threshold of acceptable human conduct.[67] Analogous standards can be applied to humanoid robot algorithms. If an algorithm's behavior surpasses the reasonable person standard—in other words, if replacing the algorithmic control with reasonable person control would not have averted the accident—this equates to fulfilling the duty of reasonable care. Consequently, in such instances, even if an accident occurs, the algorithm is not considered to have malfunctioned, and no robot liability arises. This type of algorithmic malfunction determination standard can be integrated into industry technical standards, solidifying its place in the design, development, and manufacturing stages of humanoid robots through a technical safe harbor mechanism, under which compliance with specified safety and transparency standards may serve as a defense or partial defense to liability.[68] Mirroring the "reasonable care" and "unreasonable risk of personal injury" principles in the aforementioned traffic accident liability mechanism, the humanoid robot's algorithm should be evaluated retrospectively from its behavioral effects. Algorithmic duties of care should be determined by comparison to natural person duties of care, taking into account the risk tolerance levels of different industries. If a humanoid robot algorithm fails to meet the algorithmic duty of care standards established by regulatory bodies, it must bear the corresponding risk of liability. As these "reasonable person" technical standards gain wider acceptance, the incidence of humanoid robot algorithmic malfunction and the resulting accidents they cause, which fall outside these standards, will be curtailed, and the predictability of liability will be enhanced.

III. Manufacturer as the Least Cost Avoider

The "Least Cost Avoider" principle, as articulated by Guido Calabresi in *The Costs of Accidents*, offers a compelling framework for analyzing accident liability.[69] This theory posits that, to minimize total social costs, liability should be assigned to the party best positioned to avert the risk of accidents at the lowest cost.[70] This perspective shifts the focus of accident liability mechanisms toward maximizing social welfare. From a law and economics perspective, the "Least Cost

[65]*See* Larsen v. Gen. Motors Corp., 391 F.2d 495, 504 (8th Cir. 1968).

[66]*See generally* ABBOTT, *supra* note 53.

[67]Consider the case of traffic accident law. When something goes wrong in the algorithms, they are judged based on a standard of care that largely refers to what other drivers are reasonably doing at the time.

[68]*See* Shen Weiwei, *Practices and Legal Reflections on the Technical Safe Harbor*, 4 CHINA & FOREIGN L. REV. 906, 906–22 (2023) (China).

[69]*See* GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 135, 139–58 (1970); Stephen G. Gilles, *Negligence, Strict Liability, and the Cheapest Cost-Avoider*, 78 VA. L. REV. 1291, 1292 (1992).

[70]*See id.*

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

Avoider" theory is particularly relevant when the cost of establishing legal facts is prohibitively high. In such circumstances, minimizing total social costs becomes a practically feasible and optimal objective.

In humanoid robot accidents, the excessive cost of discovering legal facts remains, under current technological and regulatory conditions, a significant challenge. Given the technical complexities inherent in humanoid robots and the potentially numerous responsible parties involved, determining liability, in no small number of cases, is both exceedingly difficult and prohibitively expensive. Applying the least cost avoider theory requires identifying, within the humanoid robot ecosystem, the participant best situated to effectively prevent such accidents at the most reasonable cost, and, post-accident, to efficiently determine and allocate responsibility. In humanoid robot accidents, the multiplicity of developers, the complexities of human-machine hybrid control, and the high attribution costs associated with machine learning algorithms render the least cost avoider approach particularly well-suited to mitigating the costs of attribution in resolving accident disputes.

The question then turns to: Who is the least cost avoider in the context of humanoid robot accidents? The relevant parties typically include the user, the owner, and the manufacturer. Each will be considered in turn.

First, the user is not the least cost avoider. Human-machine hybrid control implies that the user cannot continuously maintain complete control over all actions of the humanoid robot. Even when the user possesses some degree of control during operation, they are at an informational and technological disadvantage compared to the manufacturer. *Ex ante*, the user's control is generally limited to activation, deactivation, or setting certain automated control functions. *Ex post*, if liability is concentrated on the user, they face significant challenges in proving algorithmic malfunction and establishing the causal link between such malfunction and the resulting harm.[71] Furthermore, data from the humanoid robot at the time of the accident may be considered by the manufacturer as proprietary or potentially self-incriminating, and thus they are often reluctant to disclose it.[72] Such reluctance further complicates the discovery of algorithmic malfunctions.

Second, the humanoid robot owner is not the least cost avoider. Transferring all risk to the owner, requiring them to assume all risks associated with use at the time of purchase, would significantly disincentivize the acquisition of such high-risk products. This is particularly true given the current market competition between traditional assistive devices and humanoid robots. If owners bear all responsibility, they will likely opt for less risky traditional devices. Moreover, like the user, the owner typically has no control over the design, assembly, or quality control of the humanoid robot, and therefore lacks the capacity to mitigate the harm that may be caused by the robot's autonomous behavior. Thus, even if all liability were assigned to the owner, the owner would not be capable of effectively reducing the total social cost of accidents.

This Article argues that the manufacturer should be designated the least cost avoider.[73] First, manufacturers possess the lowest information acquisition costs regarding human-machine hybrid control and algorithmic fault, facilitating the resolution of information asymmetries in accident disputes. By designating manufacturers as the least cost avoider, the law would incentivize them to use their informational advantage to identify other culpable parties within the liability chain, employing the latter's fault as a defense to mitigate their own liability. Second, manufacturers control the technical design and modification of humanoid robots and are thus best positioned to

[71] *See generally* Tort Liability, *supra* note 57.

[72] *See, e.g.*, Brandon Vigliarolo, *Tesla Ordered to Cough Up Data for Autopilot Probe or Face Heavy Fines*, THE REGISTER (July 6, 2023, 11:58 PM), https://www.theregister.com/2023/07/06/nhtsa_tesla_crash_data.

[73] *See* Directive 2024/2853 of Oct. 23, 2024, on Liability for Defective Products and Repealing Council Directive 85/374/EEC, 2024 O.J. (L 2853) 2 (EC) (designating the manufacturer as the "least cost avoider" places China in alignment with the EU's updated Product Liability Directive, which classifies AI as a "product" and shifts the focus toward manufacturer-controlled design safety).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

manage potential accident risks. Assigning liability to manufacturers incentivizes them to minimize such risks, avoid accident costs, and provide both compensatory and even retributive remedies.[74] This is especially so for behavioral remedies, where manufacturers are uniquely positioned to implement corrective measures. Leveraging their control over both software and hardware, they can rectify humanoid robot behavior, provide compensatory relief, and even, considering the psychological dimensions of retribution, incorporate elements of retributive justice into the design itself. Moreover, even if manufacturers cannot entirely eliminate the risks associated with certain machine learning algorithms, they retain the critical option of forgoing their use to ensure product safety.[75] Third, given the current diverse landscape of humanoid robot development, manufacturers, with deeper pockets, are better positioned than individual users or owners to distribute accident costs through collectivized pricing mechanisms and insurance schemes.[76] Manufacturers occupy a central position in the development chain, thus the cost of procuring higher-quality components can be passed on to owners through increased sales prices. Similarly, insurance mechanism, such as no-fault insurance, can transfer liability risk to insurance companies and policyholders.[77] Therefore, with respect to humanoid robot accidents, the liability mechanism should designate the manufacturer as the least cost avoider, and subsequently refine the system through internal liability shifting or specific exculpatory provisions.

However, designating the manufacturer as the least cost avoider does not necessitate the imposition of strict liability. A zero-accident humanoid robot is an unrealistic expectation. While strict liability may appear superficially appealing, it could stifle manufacturers' incentives to develop and implement machine learning algorithms, ultimately harming the industry and potentially conflicting with national industrial development policies. Therefore, while identifying the humanoid robot manufacturer as the least cost avoider, regulatory agencies should, commensurate with the risk tolerance of different industries—e.g., manufacturing, healthcare, home automation, logistics—implement technology safe harbors and related liability exemptions, thereby creating a more nuanced and adaptable accident liability mechanism.[78] If fault-based liability provides sufficient incentive for manufacturers to develop humanoid robot algorithms, and can reduce user fault without increasing the risk of additional algorithmic fault, then it can generate net safety benefits for society as a whole.

[74]*See* Hu Ling, 平台视角中的人工智能法律责任 [*Legal Responsibility of Artificial Intelligence from the Perspective of Platforms*], 交 大法学 [Shanghai Jiao Tong Univ. L. Rev.] 5 (2019) (China).

[75]Sacrificing some functionalities of machine learning algorithms to achieve certainty and predictability in behavior and legal responsibility does not represent technological regression. It is crucial to recognize that conventional algorithms and machine learning algorithms are not categorically inferior or superior to one another. Different scenarios present varying technological and safety requirements. Consequently, in situations with excessively high costs of harm, applying safe and stable conventional algorithms may prove more cost-effective than utilizing cutting-edge machine learning algorithms. *See e.g.*, Andrei Aksjonov & Ville Kyrki, *A Safety-Critical Decision Making and Control Framework Combining Machine Learning and Rule-based Algorithms*, arXiv (Jan. 30, 2022), https://doi.org/10.48550/arXiv.2201.12819. *See* Bertolini, *supra* note 32 (showing similar considerations in *Artificial Intelligence and Civil Liability*).

[76]Similar considerations can be seen also in the Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, Artificial Intelligence and Civil Liability (July 2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf.

[77]*See* Omri Rachum-Twaig, *Whose Robot Is It Anyway?: Liability for Artificial-Intelligence-Based Robots*, 2020 U. Ill. L. Rev. 1141, 1164–67 (2020); Anat Lior, *Insuring AI: The Role of Insurance in Artificial Intelligence Regulation*, 35 Harv. J.L. & Tech. 467, 507 (2022).

[78]*See* Mark L. Wilde, *Railway Sparks: Technological Development and the Common Law*, 59 Am. J. Legal Hist. 444 (2019) (demonstrating that when 19th-century railways were given "statutory authority" defenses, they actually stopped innovating in safety technology because they were shielded from the cost of fires. That historical experience is instructive, but it is also importantly distinct from the contemporary humanoid robotics industry. Unlike the railways of the 1800s, today's robotics manufacturers operate in a highly competitive, reputation-sensitive global market. Firms that fail to invest in safer technologies face not legal insulation but rapid commercial displacement).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

IV. Behavioral Correction and Retribution for Humanoid Robots

As previously discussed, humanoid robots, designed to mimic human appearance, movements, and even emotional expressions, elicit complex social and emotional responses. Consequently, when a humanoid robot, exhibiting human-like behaviors and seemingly making its own choices, causes harm, the victim's experience may extend beyond the purely physical or material. The breach of trust, the sense of betrayal, and the emotional pain caused by a seemingly "sentient" being can be profound, especially when the lines of control are blurred by the hybrid control model. In such instances, traditional legal remedies, such as monetary compensation, may prove insufficient, and in particlar, China's current legal system lacks provisions for addressing these unique psychological harms. The victim may crave not only material restitution but also a symbolic acknowledgment of the robot's "wrongdoing," a form of retribution that resonates with their emotional experience of betrayal by something they perceived as having a degree of agency. Despite humanoid robots lacking human-like moral comprehension hence "no soul to damn,"[79] users might still demand an emotional form of ethical justice from them.[80] As Christina Mulligan puts it, "taking revenge against wrongdoing robots, specifically, may be necessary to create psychological satisfaction in those whom robots harm."[81]

This is where the concept of behavioral correction and retribution for humanoid robots gains relevance. While the robot itself cannot experience punishment, the act of rectifying its harmful behavior, perhaps through software updates or physical modifications, coupled with symbolic acts of "punishment," such as public deactivation, formal withdrawal from service, or even physical striking,[82] could provide a sense of closure and justice for the victim, addressing a psychological need often unmet by traditional legal remedies.

In addition, punishing robots can operate as a symbolic form of social condemnation, expressing society's disapproval of certain behaviors in ways that go beyond purely economic sanctions. As humanoid robots are entering the social sphere, they are subject to the moral expectations of that sphere. Therefore, "retribution" is not a primitive desire, but a return to the historical function of tort law: maintaining social peace and moral order.[83] Through the symbolic acts of "punishment", society conveys certain moral boundaries to the public, similar to how fines imposed on corporations not only compensate for losses but also function to publicize their violations.[84]

F. Conclusion: Towards the "Accidental Utopia"

The rapid integration of humanoid robots into society necessitates a reimagined liability mechanism to address the unique risks posed by their technological complexity—human-body simulation, artificial intelligence, and hybrid control. Traditional legal frameworks for accidents in China, rooted in fault-based liability, struggle to assign responsibility in accidents involving autonomous algorithms, emergent behaviors, and intertwined human-machine decision-making. This Article highlights three critical gaps: the inadequacy of fault standards for autonomous systems, challenges in identifying algorithmic malfunctions, and the insufficiency of traditional remedies for psychological and systemic harms.

[79]*See* Peter M. Asaro, *A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, *in* Robot Ethics: The Ethical and Social Implications of Robotics 169, 181 (Patrick Lin, Keith Abney & George A. Bekey, eds., 2012).

[80]*See* Kate Darling, *Children Beating Up Robot Inspires New Escape Maneuver System*, IEEE Spectrum (Aug. 6, 2015), https://spectrum.ieee.org/children-beating-up-robot.

[81]*See* Christina Mulligan, *Revenge Against Robots*, 69 S.C. L. Rev. 579, 580 (2018).

[82]*See* Lemley & Casey, *supra* note 39, at 1389–93 (proposing a robot death penalty).

[83]Robert J. Kaczorowski, T*he Common-Law Background of Nineteenth-Century Tort Law*, 51 Ohio St. L.J. 1127, 1173 (1990).

[84]*See* Ryan Abbott & Alex Sarch, *Punishing Artificial Intelligence: Legal Fiction or Science Fiction*, 53 U.C. Davis L. Rev. 323, 346 (2019).

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press

To mitigate these challenges, a tailored liability mechanism is proposed. Key recommendations include establishing evidentiary processes to clarify fault attribution, designating the manufacturer as the least cost avoider to incentivize risk mitigation, and integrating "reasonable person" technical standards for algorithmic behavior. Additionally, behavioral correction and symbolic retribution for robots could address victims' psychological needs.

Unless a "Goda" is developed, human physical ability will remain largely unchanged, in stark contrast to the rapidly expanding capabilities of humanoid robots. As these robots become increasingly prevalent, tasks will progressively shift from human to humanoid robot in a world that is shared by them, thereby elevating the risk of humanoid robot-related accidents. In this context, a reconstructed liability mechanism seeks to establish an "accidental utopia," where reduced accidents, fair compensation, and adaptable legal systems coexist with advanced robotic autonomy.

Acknowledgements. The author declares none.

Funding Statement. No specific funding has been declared in relation to this Article.

Competing Interests. The author declares none.

Cite this article: Shen W (2025). Reconstructing Liability Mechanism for Humanoid Robot Accidents. *German Law Journal* **26**, 1404–1422. https://doi.org/10.1017/glj.2026.10198

https://doi.org/10.1017/glj.2026.10198 Published online by Cambridge University Press