Cambridge Forum on AI: Law and Governance (2026), **2**, e19, 1–26
doi:10.1017/cfl.2026.10054

Cambridge Forum

RESEARCH ARTICLE

The liability vacuum: Why existing legal frameworks cannot govern artificial general intelligence and what must be done

Ben Chester Cheong[1,2]

[1]School of Law, Singapore University of Social Sciences, Singapore and [2]Department of Land Economy, University of Cambridge, Cambridge, UK
Email: benchestercheong@suss.edu.sg

(Received 3 March 2026 UTC; revised 21 June 2026 UTC; accepted 3 July 2026 UTC)

Abstract

Existing artificial intelligence (AI) liability scholarship focuses largely on narrow, task-specific systems. This article asks how existing liability frameworks would function – or fail – when confronted with artificial general intelligence (AGI). It examines four liability regimes – strict product liability, fault-based negligence, vicarious liability and strict liability for hazardous activities – across five jurisdictions (the European Union, United States, United Kingdom, Singapore and China), concentrating its comparative analysis on product liability, the regime most affected by digital-era reform, while situating the other three within the broader structural critique that follows. These regimes embed assumptions about predictability, human control, causal traceability and task-specificity that AGI would structurally violate. AGI challenges liability law not merely evidentially but structurally, destabilizing the core legal concepts of defectiveness, foreseeability, proximate causation and attribution of agency on which existing regimes depend. This discussion reflects the current regulatory context: the withdrawal of the EU AI Liability Directive proposal and the adoption of the revised Product Liability Directive (Directive [EU] 2024/2853). Drawing on nuclear and environmental liability analogies, the article proposes an anticipatory liability architecture: channeled strict liability with mandatory capitalization, capability-calibrated duties of care and an international compensation fund.

Keywords: artificial general intelligence; AI liability; product liability; negligence; vicarious liability; strict liability

1. Introduction: scope, contribution and the liability landscape

The global regulatory response to artificial intelligence (AI) has accelerated dramatically since 2020. The European Union has positioned itself as a first mover through the AI Act (Regulation [EU] 2024/1689), the revised Product Liability Directive (Directive [EU] 2024/2853) and the now-withdrawn proposed AI Liability Directive (COM[2022] 496 final) (European Commission, 2022). The United States has pursued a patchwork of executive orders, state-level legislation and tort litigation (Frazier, 2025; Weil, 2025). China has enacted sector-specific regulations targeting generative AI, algorithmic recommendations and deepfakes (Sheehan, 2023). Singapore has favored soft-law governance through its Model AI Governance Framework (PDPC, 2020) and AI Verify toolkit (Allen et al., 2025; IMDA, 2022). The United Kingdom has adopted a "pro-innovation" approach through sector-specific guidance, recently supplemented by the Law Commission's (2025) exploratory paper on AI and the law.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (http://creativecommons.org/licenses/by-nc-nd/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use and/or adaptation of the article.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

It is worth being precise, at the outset, about what is and is not at stake in this analysis. The AI systems that the regulatory instruments above were designed to govern are overwhelmingly instances of artificial narrow intelligence (ANI): systems trained and deployed for a specific task or bounded class of tasks – image classification, credit scoring, route optimization, language generation within a defined domain – however sophisticated their internal mechanics. Artificial general intelligence (AGI), as defined in Section 2, is a categorically different proposition: a system capable of matching or exceeding human performance across the full range of cognitive tasks, without being specifically engineered for each. No such system currently exists, and informed estimates of when, or whether, one will diverge considerably (Grace et al., 2024). The argument advanced in this article is, therefore, conditional rather than descriptive: it does not assert that AGI is imminent but argues that if and when systems approaching the AGI threshold are deployed, the four liability regimes presently relied upon to govern ANI will not transfer to them without fundamental revision. The justification for examining this question in advance of AGI's arrival is addressed directly in Section 7.1.

These regulatory instruments address liability through four distinct legal mechanisms, and this article examines all four. First, product liability holds producers liable for defective products without requiring proof of fault in the strict sense – as under the revised Product Liability Directive – and, in the United States, the Restatement (Third) of Torts: Products Liability (§§ 1–2) imposes liability on commercial sellers and distributors for harm caused by defective products, subject to distinct standards for manufacturing, design and warning defects. Second, fault-based negligence requires proof that the defendant breached a duty of care, causing harm – the regime for which the withdrawn AI Liability Directive sought to harmonize certain evidentiary rules in AI-related non-contractual civil claims and which now falls to divergent national tort laws across EU Member States (Frattone, 2025; Rimkute, 2025). Third, vicarious liability attributes to one party (an employer, a principal) liability for the acts of another (an employee, an agent) – a regime that scholars including Beckers and Teubner (2022) have proposed extending to autonomous software agents, and which Chopra and White (2011) approach through the related idea of legal agency for artificial agents. Fourth, risk-based strict liability for hazardous activities holds the operator liable regardless of fault when they engage in activities posing abnormal dangers – a regime that several scholars and the European Parliament's (2020) legislative resolution (rapporteur Voss) recommended for high-risk AI (Padovan et al., 2023; Wendehorst, 2020).

This article's central argument is that all four of these liability regimes embed assumptions about the nature of AI systems that AGI would structurally violate. The assumptions are: (a) that the system has an identifiable intended use against which its safety can be assessed (product liability); (b) that a human defendant made identifiable decisions that can be evaluated against a reasonable-care standard (negligence); (c) that the system acts within the scope of a principal-agent relationship in which the principal exercises meaningful control (vicarious liability); and (d) that the risk-creating activity is sufficiently well-defined for its dangers to be assessed and insured against (strict liability for hazardous activities). AGI, as defined in Section 2, satisfies none of these conditions.

This argument is not, however, merely theoretical. It is grounded in the current regulatory moment. The European Commission's February 2025 announcement of the withdrawal of the AI Liability Directive – officially published on 6 October 2025 (OJ C/2025/5423) – has created what Pirvan (2025) has described as a "regulatory vacuum in fault-based liability for AI." The withdrawal means that cases falling outside the scope of the revised Product Liability Directive (which covers defective products, including software, but not professional services as such, pure economic loss or nonmaterial harm beyond recognized psychological health damage) will be governed by divergent national tort laws across 27 Member States (Duffourc, 2025; Frattone, 2025). Simultaneously, the product-versus-service classification of AI systems has emerged as a central fault line in liability governance: as Pirvan (2025) observes, the strict liability regime of the revised PLD Product Liability

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

Directive applies to AI-as-product but not to AI-as-service, leaving a significant proportion of AI-caused harms to unharmonized national negligence laws. For AGI, which would resist classification as either product or service, this fault line would become an abyss.

The scholarly literature on AI liability is rich and rapidly growing. Ebers's (2021) foundational analysis of AI liability and EU consumer law (see also Ebers et al., 2009), Bertolini's (2025a, 2025b) systematic treatment of AI and civil liability, Rodríguez de Las Heras Ballell's (2025) mapping of generative AI liability scenarios under the EU framework, Buiten et al.'s ((2023)) law-and-economics analysis of AI liability and Smakman's (2025) comprehensive report on AI liability challenges in the UK have advanced our understanding considerably. The emerging literature on AI legal personhood and responsibility (Floridi & Taddeo, 2018; Novelli et al., 2025; Torrance & Tomlinson, 2025) addresses whether AI systems could bear legal responsibility directly. The Oxford Business Law Blog's analysis of post-AILD liability (Rimkute, 2025) demonstrates that EU law may continue to influence national courts even after the Directive's withdrawal. However, few existing works, to the author's knowledge, have systematically examined how each of the four principal liability regimes would fail when confronted with AGI, or proposed comprehensive anticipatory liability architecture calibrated to AGI's distinctive properties. This article aims to fill that gap.

The article proceeds in eight parts. Section 2 defines AGI for the purposes of legal analysis, drawing on the Morris et al. (2024) levels-of-AGI framework. Section 3 maps the current liability landscape across the EU, United States, United Kingdom, Singapore and China, distinguishing between product liability, fault-based liability, vicarious liability and strict liability regimes. Section 4 identifies five structural inadequacies that AGI would create across these regimes. Section 5 examines why insurance markets cannot compensate for these structural failures. Section 6 proposes a novel anticipatory liability architecture. Section 7 addresses objections. Section 8 concludes.

2. Defining AGI for legal analysis

The concept of AGI lacks a single authoritative definition, and this definitional ambiguity itself poses challenges for legal governance. In the AI safety literature, AGI is typically characterized by reference to generality, autonomy and transferability. Goertzel (2014) defines AGI as AI with the capacity to carry out virtually any cognitive task at a level equal to or exceeding that of a human being. Bostrom (2014) frames AGI in terms of its potential trajectory toward superintelligence, emphasizing the capacity for recursive self-improvement. More recently, Morris et al. (2024) have proposed a tiered framework distinguishing between "emerging AGI" (systems that match or exceed human performance across most cognitive tasks) and "full AGI" (systems that surpass the most capable humans in every cognitive domain).

For the purposes of this article, AGI is defined by four properties that are legally salient because each disrupts a specific assumption embedded in existing liability frameworks.

2.1 Domain generality

Unlike narrow AI, which is designed, trained and deployed for a specific task within a defined operational domain, an AGI system is capable of performing tasks across multiple domains without being specifically trained for each (Goertzel, 2014; Morris et al., 2024). This property disrupts the assumption of intended use that underpins product liability: if a system has no single intended use, the concept of "defectiveness" relative to its intended purpose becomes indeterminate (Rodríguez de Las Heras Ballell, 2025). The revised Product Liability Directive does not define "product" by reference to purpose (Article 4), but its defectiveness test in Article 7, read with Recital 30, assumes that the product's safety can be assessed by reference to contextual benchmarks such as its intended purpose, reasonably foreseeable use, presentation, characteristics, intended users and applicable

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

safety requirements. An AGI system deployed as a general-purpose reasoning engine resists this categorization.

2.2 Autonomous goal-setting

Current AI systems operate within objective functions defined by their developers (Cheong, 2024a). AGI, by contrast, would possess the capacity to formulate its own sub-goals in pursuit of higher-level objectives, potentially in ways that diverge from its designers' intentions (Bostrom, 2014). This property disrupts the assumption of human control that structures fault-based liability (Selbst, 2020; Wendehorst, 2020). The EU AI Act's emphasis on human oversight (Article 14) and the withdrawn proposal for an AI Liability Directive's reliance on the "duty of care" of deployers presuppose that a human actor can meaningfully supervise and correct the system's operations. When a system can set its own goals, the locus of decision-making shifts from the human operator to the system itself, creating a gap in the attribution of responsibility (Beckers & Teubner, 2022).

2.3 Causal opacity

While the "black box" problem is already recognized in relation to deep learning systems, more general and agentic AI systems may intensify this opacity. A narrow AI system producing an erroneous medical diagnosis operates within a bounded task environment, even if its internal model remains difficult to interpret. By contrast, an AGI-like system causing harm through a chain of reasoning spanning multiple domains – for example, combining medical knowledge, economic incentives and strategic considerations – could generate causal pathways that are not only technically opaque but practically resistant to reconstruction by human investigators (Faroldi, 2025). This would strain conventional doctrines of factual and proximate causation. It also exposes the limits of the now-withdrawn AI Liability Directive: Article 3's disclosure and evidence-preservation mechanism and Article 4's rebuttable presumption of causality were designed to alleviate evidential asymmetry in claims involving AI systems but may be inadequate where the legally relevant causal chain is diffuse, multi-domain and only partially reconstructable (Madiega, 2023).

2.4 Self-modification

Perhaps most significantly, AGI systems may possess the capacity to modify their own architecture, parameters or training data. A system that has been materially altered since its initial deployment is, in a meaningful sense, no longer the same product that was placed on the market. This disrupts the temporal framework of product liability, which assumes that the product's characteristics at the time of placing on the market can be assessed and that the producer's obligations are defined by reference to that moment. The revised Product Liability Directive's extension to software updates and upgrades (Article 11(2)) addresses post-market modifications within the manufacturer's control, but not autonomous self-transformation outside such control.

These four properties are not speculative extrapolations but direct implications of the capabilities that AGI developers are actively pursuing (Bubeck et al., 2023; Grace et al., 2024; Salmon et al., 2023). They collectively suggest that AGI would occupy a fundamentally different legal category from the AI systems currently addressed by regulation – one that existing frameworks are not designed to accommodate (Chesterman, 2021; Turner, 2018).

3. The current global liability landscape: implicit assumptions and structural limitations

This section maps the principal liability frameworks applicable to AI across five jurisdictions, identifying the assumptions about AI capability that each framework embeds (Giuffrida, 2019; Lemley & Casey, 2019; Soyer & Tettenborn, 2022). Because product liability reform has been the focus of recent

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

legislative activity in most of the jurisdictions surveyed, the discussion below concentrates primarily on that regime; negligence, vicarious liability and strict liability for hazardous activities are addressed comparatively rather than exhaustively and are revisited in the structural analysis of Section 4, which applies to all four regimes equally.

3.1 European Union

The EU's approach to AI liability rests on three pillars, each now in a different state of development. The AI Act (Regulation [EU] 2024/1689), which entered into force on 1 August 2024, establishes a risk-based regulatory framework but does not itself create liability rules. The revised Product Liability Directive (Directive [EU] 2024/2853), adopted on 23 October 2024 and requiring transposition by 9 December 2026, modernizes the 1985 regime to encompass software and AI systems. The proposed AI Liability Directive, which would have introduced fault-based evidentiary mechanisms including rebuttable presumptions of causation, was announced for withdrawal by the European Commission in February 2025 and formally withdrawn in October 2025 (European Commission, 2025).

The withdrawal of the AI Liability Directive is particularly significant for the analysis of AGI liability. The Directive was designed, in part, to address a specific problem: the difficulty that claimants face in proving fault and causation when AI systems are opaque and complex (Ebers, 2021). Its proposed mechanisms – including rights of access to evidence (Article 3) and a rebuttable presumption of a causal link between fault and AI output (Article 4) – assumed that a relevant defendant (provider, deployer or authorized representative) could be identified and that the causal relationship between their failure to comply with a duty of care and the resulting harm could, with appropriate procedural assistance, be established.

These assumptions hold for narrow AI. A deployer who fails to implement the human oversight measures required by the AI Act may be presumed to have caused the harm that results from the unsupervised system's output. But AGI disrupts this logic. If the system's harmful action results from autonomous reasoning across domains, the deployer's failure to supervise is not the cause of the specific harm – it is the cause of the general condition (the system's unsupervised operation) within which the system's own autonomous decision was the proximate cause. The causal gap between human fault and AI-generated harm widens to the point where the presumption of causation becomes a legal fiction.

The revised Product Liability Directive fares somewhat better but remains structurally limited (Ebers, 2021; Wagner, 2023; Wendehorst, 2020). Its expanded definition of "product" to include software (Article 4, point (1)), its recognition that products can continue to learn or acquire new features after being placed on the market or put into service (Article 7(2), point (c)), and its extension of liability to defective components, including intangible components and related services integrated into or inter-connected with products (Articles 4, points (3)–(4), and 8(1), point (b)), represent important advances. However, the Directive's concept of "defectiveness" remains tethered to the safety that a person is entitled to expect, assessed objectively by reference to the safety that the public at large is entitled to expect (Article 7(1); Recital 30), which presupposes that a product's behavior is sufficiently predictable for such expectations to be formed (Bertolini, 2025a, 2025b). An AGI system's behavior may be unpredictable not because of a defect but because unpredictability is an inherent feature of general intelligence.

3.2 United States

The United States lacks a comprehensive federal framework for AI liability, relying instead on existing tort doctrines (negligence, strict liability, products liability under the Restatement [Third] of Torts), sectoral regulation (the Federal Trade Commission Act, the Fair Housing Act and state consumer

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

protection statutes) and an evolving body of litigation (Calo, 2015; Lemley & Casey, 2019; Rachumtwaig, 2020). The Character.AI litigation (*Garcia v. Character Technologies, Inc.*, No. 6:24-cv-01903 [M.D. Fla.]), in which the parents of a teenager who died by suicide alleged that an AI chatbot contributed to his death through wrongful death, product liability and negligence claims, exemplifies the strain that even narrow AI places on existing tort categories. A May 2025 ruling by Judge Conway allowing most claims to proceed – rejecting the defendant's argument that chatbot output constitutes protected First Amendment speech – marked an important early recognition that AI chatbots may be treated as products subject to safety standards rather than simply as expressive media (Frazier, 2025). The case subsequently settled in January 2026, amid congressional scrutiny, an FTC Federal Trade Commission investigation and state legislative activity concerning AI chatbots and child safety.

The rapid proliferation of state-level AI legislation in the United States has added a further layer of complexity. Colorado's AI Act, enacted in May 2024, represented the most comprehensive state-level intervention to date, focusing on algorithmic discrimination in high-risk decisions affecting consumers. Texas enacted the Responsible Artificial Intelligence Governance Act in June 2025. California's CCPA California Comsumer Privacy Act regulations on automated decision-making technology, effective January 2027, will impose specific disclosure and opt-out requirements for AI-driven significant decisions. Utah has enacted legislation specifically targeting AI-powered mental health chatbots. According to some estimates, over 1,000 AI-related bills were introduced across state legislatures during the 2025 legislative session (Gluck, 2025; Harris, 2025; MultiState, 2026). This legislative hyperactivity creates both opportunities and risks: it generates innovation in regulatory approaches but also produces the jurisdictional fragmentation that is particularly problematic for AGI systems that would operate across state and national boundaries.

The U.S. Senate Judiciary Committee's September 2025 hearings on AI chatbot harms, followed by the introduction of the bipartisan AI LEAD Act by Senators Hawley and Durbin, signal a potential shift toward federal engagement with AI liability (US Senate Committee on the Judiciary, 2025a, 2025b). However, the legislation focuses on applying product liability concepts to AI systems, rather than addressing the broader structural challenges that AGI would present. The conceptual framework of U.S. product liability law – rooted in the Restatement (Third)'s distinction between manufacturing defect, design defect and failure to warn – remains fundamentally oriented toward physical products with identifiable design choices (American Law Institute, 1998).

Two further lines of litigation illustrate how AI liability challenges are proliferating across sectors. In *Mobley v. Workday, Inc.*, 740 F. Supp. 3d 796 (N.D. Cal. 2024), Judge Lin held that an AI vendor could face direct liability for employment discrimination under federal anti-discrimination laws when its algorithmic screening tools participate in hiring decisions. The court reasoned that Workday's AI software was not simply implementing employer criteria but was itself "participating in the decision-making process," acting as an agent whose biases could ground a cognizable discrimination claim. The case was subsequently preliminarily certified as a collective action in 2025, potentially encompassing millions of applicants (Morris, 2025). Separately, in *Huskey v. State Farm Fire & Casualty Co.*, No. 22 C 7014 (N.D. Ill, 11 September 2023), the court allowed a disparate-impact claim under § 3604(b) of the Fair Housing Act to proceed where the plaintiffs alleged that State Farm's algorithmic fraud-detection tools systematically subjected Black policyholders to greater scrutiny than white policyholders (Karvosky, 2025). These cases demonstrate that the attribution and bias challenges already arising from narrow AI systems – who is liable when an algorithm discriminates? – would be vastly amplified by AGI's cross-domain autonomy.

For AGI, the U.S. framework faces three distinctive challenges. First, the Restatement (Third)'s categories of manufacturing defect, design defect and failure to warn (Sections 1 and 2) assume that the product's defect can be evaluated by reference to an identifiable departure from intended design, a reasonable alternative design or reasonable instructions or warnings. An AGI system whose capabilities emerge from training on vast datasets and whose behavior is shaped by autonomous learning

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

does not lend itself to this analysis. The emergent nature of AGI capabilities – where system behavior arises from the interaction of billions of parameters rather than from discrete design decisions – means that the Restatement's requirement, in design-defect cases, to identify a "reasonable alternative design" becomes practically impossible. What would a reasonably safe alternative design for general intelligence look like? The question has no meaningful answer within the existing doctrinal framework.

Second, the foreseeability requirement in negligence law – the foundational *Palsgraf* principle (*Palsgraf v. Long Island Railroad Co.*, 248 N.Y. 339) – assumes that a defendant's duty of care extends to harms that are reasonably foreseeable. When an AGI system acts across domains in novel ways, the category of foreseeable harms becomes effectively unbounded, collapsing the distinction between duty and no-duty that structures negligence analysis. The autonomous vehicle litigation to date illustrates this challenge in embryonic form: in the Uber ATG fatality in Tempe, Arizona (2018), where the automated driving system failed correctly to classify and respond to a pedestrian, and in the Florida Tesla Autopilot verdict of 2025, where a jury awarded $243 million after finding that Tesla bore partial responsibility for a fatal crash involving Autopilot's design and representations about its capabilities, courts have struggled to apply traditional foreseeability doctrine to AI systems whose decision-making processes resist human comprehension (DeArman, 2019; Stempel & Roy, 2025). AGI would represent not merely a quantitative expansion of the range of foreseeable harms but a qualitative transformation: the very concept of foreseeability presupposes a degree of predictability in the system's behavior that AGI would lack by design.

Third, the absence of a comprehensive federal AI liability framework means that any AGI-specific liability regime would need to be constructed from the ground up, without the institutional infrastructure that the EU AI Act provides. The Rhode Island Senate Bill 358, which would impose strict liability on AI developers when AI systems engage in conduct that would be wrongful if performed by a human, represents one attempt to close this gap (Weil, 2025). The bill is designed to address situations where an AI system autonomously engages in harmful conduct (such as sending phishing emails), and no current legal doctrine holds any party liable. While narrow in scope, the Rhode Island bill illustrates the kind of targeted strict liability mechanism that could form one component of a broader AGI framework.

President Trump's December 2025 Executive Order directing the establishment of an AI Litigation Task Force to challenge state AI laws further complicates the picture, signaling a federal preference for minimal regulation that sits in tension with the need for robust AGI liability rules (The White House, Ensuring a National Policy Framework for Artificial Intelligence, 2025). The tension between federal deregulatory impulses and state-level regulatory activism creates a deeply uncertain legal environment that would be particularly problematic for AGI deployment, where legal certainty is essential for both safety assurance and investment confidence.

3.3 United Kingdom

The United Kingdom has adopted a "pro-innovation" approach to AI regulation, eschewing comprehensive legislation in favor of sector-specific guidance delivered through existing regulators (UK Government, 2023). The Law Commission's July 2025 paper *AI and the Law* marks a significant intervention, explicitly acknowledging the possibility of "liability gaps" where no natural or legal person is liable for harms caused by AI systems and raising the "potentially radical option" of granting some form of legal personality to AI systems (Cheong, 2024b; Law Commission, 2025). This is a significant official UK acknowledgment that existing frameworks may be structurally inadequate for highly autonomous AI.

However, the Law Commission's discussion remains exploratory. The UK's current position relies on the Consumer Protection Act 1987 (implementing the original Product Liability Directive), the common law of negligence and emerging sector-specific guidance. None of these frameworks

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

addresses the distinctive properties of AGI. The UK's reliance on common law evolution as the primary mechanism for legal adaptation is particularly ill-suited to AGI, which would present challenges too fundamental and too urgent for incremental case-by-case development. The common law's strength lies in its ability to adapt incrementally to novel factual situations within an established conceptual framework; AGI would require not incremental adaptation but wholesale re-conceptualization of the framework itself.

The UK's post-Brexit regulatory autonomy provides both an opportunity and a risk in this context (Cheong, 2022b). On the one hand, the UK is no longer bound by the EU's regulatory choices and could, in principle, develop a more innovative and forward-looking AGI liability framework. On the other hand, the UK's "pro-innovation" regulatory philosophy, which explicitly prioritizes avoiding "legislating too early" for AI, creates institutional resistance to anticipatory regulation. Recent developments reinforce this growth-oriented trajectory: the Data (Use and Access) Act 2025 adopts more permissive rules on automated decision-making, while initiatives such as the Digital Standards Strategy and designated "AI Growth Zones" focus on development, infrastructure and international standards influence. At the same time, the Digital Regulation Cooperation Forum's work on agentic AI shows that UK regulators are beginning to confront regulatory uncertainties and accountability challenges in more autonomous systems (Department for Science, Innovation and Technology, 2025a, 2025b, 2025c, 2026; Digital Regulation Cooperation Forum, 2025).

The tension between these impulses – the opportunity for regulatory innovation and the preference for regulatory restraint – remains unresolved.

3.4 Singapore

Singapore's approach to AI governance has been characterized by pragmatic soft law. The Model AI Governance Framework (2019, updated 2020) provides nonbinding guidance on transparency, accountability and human-centric AI deployment. The Infocomm Media Development Authority's AI Verify framework (2022) offers a testing toolkit for responsible AI deployment. Singapore subsequently developed the Model AI Governance Framework for Generative AI (IMDA, 2024, 2026a) and, in 2026, launched a Model AI Governance Framework for Agentic AI, which provides guidance on responsible deployment of AI agents while emphasizing that humans remain ultimately accountable (IMDA, 2026b). Alongside these soft-law instruments, Singapore has updated its National AI Strategy, proposed MAS guidelines on AI risk management for financial institutions and expanded AI testing and assurance initiatives (IMDA, 2026c; MAS, 2025, 2026; MDDI, 2026; Smart Nation Singapore, 2026). These developments confirm Singapore's pragmatic, innovation-friendly approach but also show increasing attention to the governance challenges posed by more autonomous and agentic AI systems.

The city-state's approach reflects a deliberate choice to avoid premature regulation that might stifle innovation in an economy that depends heavily on technology-sector growth (Allen et al., 2025).

For liability specifically, Singapore relies on its general tort law framework, the Consumer Protection (Fair Trading) Act and sector-specific regulation (for example, the Financial Advisers Act for AI-assisted financial services). While there is no dedicated AI liability statute, Singapore's approach reflects a deliberate and considered regulatory strategy. The Model AI Governance Framework's emphasis on transparency, accountability and human-centric design (PDPC, 2020) has been internationally recognized as a leading soft-law framework, and Singapore's pragmatic, innovation-friendly approach has enabled it to attract significant AI investment and talent as part of the Smart Nation Initiative. As Cheong has argued, transparency and accountability are foundational to responsible AI governance, and Singapore's frameworks provide a strong basis upon which more specific liability provisions could be built as AI systems become more capable (Cheong, 2024a). Singapore's institutional strengths – its strong rule of law, sophisticated regulatory infrastructure and strategic position as a hub

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

for technology companies operating across ASEAN Association of Southeast Asian Nations – position it well to develop anticipatory liability frameworks calibrated to AGI's distinctive challenges.

3.5 China

China has adopted one of the most interventionist approaches to AI regulation among major jurisdictions, but its interventions have been targeted at specific applications rather than general liability principles. The Interim Measures for the Management of Generative AI Services (CAC, 2023), the Provisions on the Administration of Algorithmic Recommendation in Internet Information Services (CAC, 2021) and the Provisions on the Administration of Deep Synthesis in Internet-based Information Services (CAC, 2022; Library of Congress, 2023) regulate specific AI capabilities within China's broader framework of internet content control. Since 2025, China has further expanded this framework through mandatory rules on the labeling of AI-generated synthetic content (CAC, 2025), AI safety governance frameworks (Cade Project, 2025) and standardization initiatives (Reuters, 2024), while 2026 legislative plans indicate movement toward more comprehensive AI legislation (Zhang, 2026). China's Civil Code, which entered into force in 2021, provides general tort liability rules, including provisions on product liability (Articles 1202–1207) and liability for damage caused by ultra-hazardous activities (Articles 1236–1244) (NPC, 2020).

China's framework is notable for its emphasis on state control over AI development and deployment, reflecting a governance philosophy that prioritizes social stability and state authority. This approach may provide a stronger institutional basis for AGI regulation than Western liberal frameworks, as the state retains broad discretionary power to intervene in technology markets (Carvalho, 2026). However, China's sector-specific approach shares the same limitation as other jurisdictions: its regulations are designed for specific types of AI systems and do not address the cross-domain, autonomous properties of AGI.

4. Five structural inadequacies: why current frameworks cannot govern AGI

Drawing on the comparative analysis above, this section identifies five structural inadequacies that would render existing liability frameworks incoherent if applied to AGI.

4.1 The defectiveness paradox

Product liability, whether under the EU's revised Product Liability Directive, the U.S. Restatement (Third) or national analogues, requires a finding that the product was "defective" – that is, that it deviated from the level of safety that the relevant standard demands. Under the revised PLD, a product is defective if it "does not provide the safety that a person is entitled to expect or that is required under Union or national law" (Article 7(1)). Under the Restatement (Third), a design defect exists where "the foreseeable risks of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design" and the omission of that alternative design renders the product not reasonably safe (Section 2(b)).

AGI creates a paradox within both formulations (Buiten et al., (2023); Faroldi, 2025; Selbst & Barocas, 2018). A system designed for general intelligence is designed to be unpredictable in its specific outputs – this is a feature, not a defect, of generality. The capacity to generate novel solutions, to combine knowledge across domains and to adapt to unforeseen situations is precisely what distinguishes AGI from narrow AI (Faroldi, 2025; Goertzel, 2014; Morris et al., 2024). If the system produces harmful outputs through a process of autonomous reasoning that its developers did not specifically anticipate, is the system defective? Or is it functioning exactly as designed?

The paradox is most acute in the EU context. The revised PLD's list of factors relevant to assessing defectiveness includes the product's "ability to continue to learn or acquire new features" after

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

being placed on the market (Article 7(2)(c)) and its "reasonably foreseeable use" (Article 7(2)(b)). For AGI, these factors generate circular reasoning: the effect of continued learning is the production of novel, unpredictable outputs and the reasonably foreseeable use of a general-purpose system is, by definition, any use. The defectiveness inquiry collapses into a question about whether general intelligence itself is defective.

Wagner (2019) and Bertolini (2025a, 2025b) have recognized versions of this problem in the context of autonomous and learning AI systems, but AGI intensifies it qualitatively. A narrow AI system that learns in unexpected ways can be compared to an alternative design that restricts learning. An AGI system cannot be redesigned to be less general without ceasing to be AGI. The defectiveness framework requires comparison to a safer alternative, but the only alternative to general intelligence is narrow intelligence – a fundamentally different product.

A further dimension of the defectiveness paradox emerges once the remaining circumstances listed in Article 7(2) are read together with the AI Act's own safety architecture. Article 7(2)(d) directs courts to consider the "reasonably foreseeable effect on the product of other products that can be expected to be used together with [it], including by means of inter-connection" – a factor designed for products with a bounded and identifiable set of compatible components. An AGI system capable of autonomously selecting and combining external tools, data sources or other AI systems in pursuit of a goal has no such bounded set of interconnections: the class of "products reasonably expected to be used together with it" is, in principle, unbounded and self-determined by the system rather than specified by its designer, compounding the indeterminacy already identified in relation to the continued-learning factor of Article 7(2)(c). Article 7(2)(f) and (g) deepen the difficulty by tying the defectiveness assessment to compliance with applicable product-safety requirements – including, for AI systems, the AI Act's risk-management (Article 9) and human-oversight (Article 14) obligations – and to the adequacy of any subsequent product recalls or regulatory interventions. Both provisions presuppose that compliance can be assessed against a fixed, enumerable set of risks identified at the point of deployment, and that "the product" targeted by a recall remains stable enough to be withdrawn or corrected.

For an AGI system whose risk profile evolves through autonomous learning, and whose self-modification may render the deployed system materially different from the one originally assessed for AI Act compliance, satisfying Articles 9 and 14 at the point of deployment offers little assurance about the system's defectiveness at the point of harm, and a "recall" may not meaningfully capture or reverse the relevant capability (Faroldi, 2025). The interaction between Article 7(2) and the AI Act's safety obligations, therefore, does not resolve the defectiveness paradox; it relocates the paradox from the product itself to the regulatory compliance regime meant to certify the product's safety in the first place.

4.2 The causation abyss

Causation is the conceptual backbone of liability law, connecting the defendant's conduct (or the product's defect) to the claimant's harm. In both fault-based and strict liability regimes, the claimant must establish a causal link between the legally relevant factor and the damage suffered (Steel, 2024). The black box problem of deep learning systems is well-documented: the internal processes by which a neural network produces an output may resist human comprehension (Bathaee, 2018; Selbst & Barocas, 2018). Existing frameworks have attempted to address this through evidentiary mechanisms: the withdrawn AI Liability Directive proposed a rebuttable presumption of causation (Article 4), and the revised PLD allows courts to presume defectiveness, causation or both where, despite disclosure of evidence, the claimant faces "excessive difficulties" in proving defectiveness or causation due to "technical or scientific complexity" and demonstrates that defectiveness or causation is likely (Article 10(4)).

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

AGI does not merely present evidentiary difficulties in establishing causation; it threatens to dissolve the concept itself (McLean et al., 2023; Padovan et al., 2023; Selbst, 2020). When an AGI system causes harm through a chain of autonomous reasoning spanning multiple domains – for example, by combining medical information, financial incentives and legal knowledge to advise a patient to discontinue treatment in a way that maximizes insurance savings – the causal chain between the system's design or the deployer's supervision and the specific harm becomes attenuated to the point of meaninglessness. The legally relevant cause of the harm is the system's own autonomous reasoning, not any identifiable defect or failure of oversight.

This creates what might be called a "causation abyss" – a gap between the human actors who created and deployed the system and the system's own autonomous actions that no procedural mechanism can bridge (Matthias, 2004; Da Silva, 2024). Rebuttable presumptions of causation work when the gap is evidentiary: the causal link exists but is difficult to prove. They do not work when the gap is deeper than evidential: the connection between human conduct and AI-generated harm may no longer exist in the form that liability law requires for meaningful attribution of responsibility (Oimann & Tollon, 2025; Santoni de Sio & Mecacci, 2021).

4.3 The attribution crisis

Existing AI liability frameworks distribute responsibility across the "AI value chain": developers, providers, deployers, distributors, importers and authorized representatives each bear obligations calibrated to their role (Beckers & Teubner, 2023; European Expert Group on Liability and New Technologies, 2019). The EU AI Act and the revised PLD both structure their obligations around these roles. This distribution assumes that each actor in the chain makes identifiable decisions that contribute to the risk of harm, and that responsibility can be allocated proportionally (Kiener, 2025).

AGI disrupts this distribution by introducing an actor that is not anticipated by the framework: the system itself (Beckers & Teubner, 2022; Chopra & White, 2011). When an AGI system autonomously decides to apply a technique learned in one domain to a problem in another, the decision to act in this way is not attributable to the developer (who did not program the specific cross-domain transfer), the deployer (who may not have anticipated the specific application) or the user (who may have given only a general instruction). The system's autonomous agency creates a gap in the chain of attribution that cannot be filled by assigning vicarious liability to any human actor (Novelli et al., 2025; Wagner, 2019).

The European Parliament's (2017) proposal for "electronic personhood" was an attempt to address this gap by granting legal personality to autonomous systems, allowing them to bear liability directly (Cheong, 2024b; European Parliament, 2017). The proposal was widely criticized – including by a group of 285 AI and robotics experts in an Open Letter to the European Commission: Artificial Intelligence and Robotics (2018) and by Floridi and Taddeo (2018) – on the grounds that it would risk misplacing moral responsibility, causal accountability and legal liability, allowing human actors to shield themselves behind the AI's nominal personality. The criticism was valid for narrow AI, where human actors retain meaningful control. For AGI, the argument is more finely balanced: if the system genuinely acts autonomously, attributing liability to humans who did not and could not have directed its actions may be equally fictitious.

4.4 The temporal dislocation

Product liability frameworks anchor the producer's obligations to a specific temporal moment: the time at which the product was "placed on the market" or "put into service." The revised PLD extends this by requiring account to be taken of the product's ability to continue to learn or acquire new features after that moment (Article 7(2)(c)), and by preserving liability for defectiveness caused by

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

software updates or upgrades, the lack of safety-preserving software updates or upgrades or substantial modifications within the manufacturer's control (Article 11(2)). The AI Act imposes ongoing post-market surveillance obligations. These extensions address the problem of evolving software but retain the assumption that the product's evolution is directed by, or remains within the control of, identifiable human or legal actors.

AGI's capacity for self-modification severs the temporal link between the product placed on the market and the product as it exists at the time of harm (Pirvan, 2025; Rodríguez de Las Heras Ballell, 2025). An AGI system that has autonomously modified its own parameters, reasoning strategies or knowledge base may be, in functionally relevant respects, a different system from the one that was originally deployed (Faroldi, 2025). The question of which entity – the original product or the modified system – caused the harm becomes legally undecidable. If the original product is the basis for liability, the producer is held responsible for a system it did not create. If the modified system is the basis, there is no identifiable producer to hold liable.

This temporal dislocation is analogous to the challenge the present author has identified in the context of digital legal personality, where the disconnect between the legal identity of an avatar or AI system and the real person behind it creates accountability gaps that existing veil-piercing doctrines struggle to address (Cheong, 2022a). The problem of attributing legal responsibility when the entity causing harm is structurally distinct from its human creators – whether through corporate form, avatar intermediation or autonomous AI operation – is a recurring challenge in technology law (Cheong, 2024b; Papastergiou et al., 2026). As the present author has argued elsewhere, transparency and accountability mechanisms are essential to any governance framework for AI systems, yet existing legal and regulatory structures remain inadequate to address the opacity and autonomy of advanced AI (Cheong, 2024a). In both contexts, the law's reliance on temporal and conceptual anchoring creates a structural gap that permits harm without accountability.

4.5 The jurisdictional fragmentation

Finally, AGI exposes a problem that is already acute for narrow AI but would become intractable for general-purpose systems: the fragmentation of liability regimes across jurisdictions. An AGI system developed in one jurisdiction, trained on data from multiple jurisdictions, deployed by a company in a third jurisdiction and accessed by users worldwide generates a web of potential liability that no single national framework can comprehensively address (Perboli et al., 2025). The EU's approach to this problem – asserting jurisdiction over AI systems that produce effects within the EU regardless of where they are developed – provides a partial solution but cannot address harms that cross multiple jurisdictions or that occur in jurisdictions without robust AI liability regimes.

The withdrawal of the AI Liability Directive exacerbates this problem within the EU itself. Without harmonized fault-based liability rules, Member States will rely on their diverse national tort laws, creating the very fragmentation that the Directive was designed to prevent (Duffourc, 2025; Nannini, 2025). For AGI, which would operate across borders as a matter of course, the resulting patchwork of national rules would create both under-protection (where victims cannot access effective remedies) and over-deterrence (where developers face unpredictable liability across multiple jurisdictions) (Perboli et al., 2025).

The convergence of these five structural inadequacies produces a compound failure that is greater than the sum of its parts. In isolation, each inadequacy might be addressed through targeted reform: the defectiveness paradox through revised safety standards, the causation abyss through enhanced evidentiary presumptions and the attribution crisis through expanded vicarious liability. But together, they reveal a fundamental misalignment between the conceptual architecture of liability law and the nature of AGI as an artifact. Liability law presupposes identifiable human decisions, traceable causal chains, assessable product characteristics, temporal stability and jurisdictional boundaries. AGI, as

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

defined in Section 2, possesses none of these properties (Faroldi, 2025). The conclusion is not that liability law needs reform but that it needs re-conceptualization.

It is worth emphasizing that these inadequacies are not merely theoretical. They are already observable in embryonic form in litigation involving current AI systems. The *Garcia v. Character Technologies* litigation illustrates the causation abyss: the plaintiffs alleged a causal link between specific design choices and the chatbot's allegedly harmful interactions with a vulnerable teenager, but the case settled in January 2026 before the court could resolve the underlying doctrinal questions. The *Mobley v. Workday* litigation illustrates the AI attribution crisis: the court allowed claims to proceed on the basis that Workday could be treated as an agent of the deploying employer, but the complex question of how to legally apportion liability between an AI vendor and the deploying employer remains unresolved (*Mobley v. Workday, Inc.*, 740 F. Supp. 3d 796 [N.D. Cal. 2024]). In *Huskey v. State Farm* (N.D. Ill. 2023), the plaintiffs alleged that machine-learning fraud-detection algorithms systematically discriminated against Black homeowners, illustrating how algorithmic opacity creates evidentiary challenges that existing disparate-impact doctrine was not designed to address. The autonomous vehicle cases – from the Uber ATG pedestrian fatality in Arizona (2018), where the automated driving system failed correctly to classify and respond to a pedestrian, to the 2025 Florida Tesla Autopilot verdict ($243 million), involving a fatal collision with people near a parked vehicle – illustrate the defectiveness paradox: is an autonomous system that misclassifies a pedestrian or fails to detect a stationary hazard "defective," or is it operating within the inherent limitations of probabilistic machine perception? AGI would amplify each of these challenges to the point where existing frameworks cease to function as effective mechanisms for compensation and deterrence.

5. The insurance gap: why market mechanisms cannot compensate for structural legal failure

A common response to perceived gaps in liability frameworks is to rely on insurance markets or fund-based compensation mechanisms to allocate risk efficiently (Erdélyi & Erdélyi, 2020; Yoshikawa, 2019). The logic is straightforward: if liability rules create uncertainty, insurers or compensation funds can price or pool that uncertainty, and the cost of insurance or contribution to a fund serves as both a compensation mechanism and a deterrent against excessive risk-taking. This market-based approach has worked reasonably well for many categories of product liability, where historical claims data enable actuarial modeling and the range of potential harms is bounded (Soyer & Tettenborn, 2022).

AGI fundamentally undermines this logic (Buiten et al., (2023); Smakman, 2025). Insurance pricing depends on the ability to estimate the probability and magnitude of potential losses. For AGI, neither quantity is estimable with the precision that actuarial modeling requires. The probability of harm from an AGI system depends on the system's autonomous decisions, which are by definition unpredictable (Bostrom, 2014). The magnitude of potential harm is effectively unbounded: an AGI system deployed in critical infrastructure, financial markets or healthcare could cause harm ranging from individual injury to systemic disruption. The catastrophe bond model proposed by some commentators (Weil, 2025) addresses the magnitude problem but not the probability problem.

Furthermore, the moral hazard concerns that attend any liability insurance regime are amplified for AGI. If developers and deployers know that insurance will cover the costs of AGI-caused harm, their incentives to invest in safety are correspondingly reduced. This moral hazard is more severe for AGI than for conventional products because the safety investments required – including containment protocols, ongoing monitoring and alignment research – are extraordinarily expensive and their effectiveness is uncertain (Trout, 2026). The insurance market's standard tool for managing moral hazard – deductibles, coinsurance requirements and experience-rated premiums – cannot function effectively when the insurer cannot assess the insured's risk management practices with confidence.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

These considerations suggest that insurance should be one component of an AGI liability regime but cannot serve as its foundation. The primary mechanism must be regulatory, establishing mandatory safety standards and operational requirements that reduce the probability of harm before it occurs. Insurance then serves as a secondary mechanism, providing compensation for residual risk that regulation cannot eliminate. This layered approach – regulation for prevention, capitalization and insurance for compensation – is precisely what the nuclear and environmental liability conventions have adopted, and it provides the most promising model for AGI (Trout, 2024).

6. Toward an anticipatory liability architecture for AGI

The structural inadequacies identified above are not problems that can be resolved by incremental reform of existing frameworks (Duffourc & Gerke, 2023; Soyer & Tettenborn, 2022). They require a fundamentally new approach: an anticipatory liability architecture that is designed for the distinctive properties of AGI before such systems are deployed, rather than constructed reactively after harm has occurred (Guston, 2014; Stilgoe et al., 2013).

The concept of anticipatory governance has been developed in the science and technology studies literature (Guston, 2014; Stilgoe et al., 2013) and has been applied to emerging technologies including nanotechnology and gene editing. Its application to AGI liability has not, to the author's knowledge, been systematically developed (Mueller, 2025). This section proposes three components of such an architecture, drawing on analogies from existing international liability regimes.

6.1 Mandatory capitalization reserves: learning from nuclear liability

The international nuclear liability regime, established by the Paris Convention (1960) and the Vienna Convention (1963) and supplemented by the Convention on Supplementary Compensation for Nuclear Damage (1997), addresses a challenge that shares structural features with AGI liability: the potential for catastrophic harm that exceeds any individual actor's capacity to compensate, combined with the difficulty of attributing fault in complex technological systems (Dominguez & McCauley, 2025; Trout, 2024). Nuclear liability conventions address these challenges through three mechanisms: channeling liability to the operator, imposing strict (no-fault) liability and requiring mandatory financial security (insurance or other financial guarantees) up to defined limits (Schwartz, 2006).

An AGI liability regime could adapt these principles (Erdélyi & Erdélyi, 2020; Rachum-twaig, 2020; Yoshikawa, 2019). First, liability for AGI-caused harm should be channeled to a single identifiable entity – the "AGI operator," defined as the entity that deploys the AGI system for public-facing or commercially significant purposes. This usage is deliberately narrower than the EU AI Act's definition of "operator" (Article 3(8)), which is an umbrella term covering the provider, product manufacturer, deployer, authorized representative, importer and distributor collectively. The AGI operator proposed here corresponds most closely to the Act's "deployer" (Article 3(4)) or, where the deploying entity is also the system's developer, its "provider" (Article 3(3)). The term "operator" is retained nonetheless, in preference to either "deployer" or "provider," because it tracks the channeling function performed by the equivalent term in the nuclear liability conventions discussed above, and because it designates a liability-bearing role rather than a regulatory-compliance role. This avoids the attribution crisis by designating a single responsible party regardless of the causal complexity of the system's actions (Beckers & Teubner, 2023). Second, liability should be strict: the AGI operator should be liable for harm caused by the system's operations without requiring proof of fault, defect or foreseeability (Wendehorst, 2020). This addresses both the defectiveness paradox and the causation abyss by removing the need to trace the harm to a specific defect or human failure. Third, the AGI operator should be required to maintain mandatory capitalization reserves – ring-fenced financial assets – sufficient to compensate foreseeable harms, with the level of required capitalization calibrated to the system's capabilities and deployment context.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

This approach has the additional advantage of internalizing the costs of AGI risk. Under current frameworks, the costs of AI-caused harm are largely externalized to victims and the public, creating a subsidy for AI development that distorts investment decisions (Buiten et al., (2023); Buolamwini & Gebru, 2018). Mandatory capitalization would ensure that the expected costs of harm are reflected in the economics of AGI development and deployment.

The capitalization requirement should be designed with sensitivity to the distributional effects of AGI development. Currently, AGI research is concentrated in a small number of large technology companies and well-funded laboratories, predominantly in the United States (Negele et al., 2025). Excessive capitalization requirements could entrench this concentration by creating barriers to entry for smaller developers and developers in less wealthy jurisdictions. The regime should therefore include tiered requirements: higher capitalization for systems deployed at scale in high-risk domains, lower requirements for research-stage systems, and exemptions or subsidies for academic research and public-interest development. The nuclear liability regime's approach – which sets minimum liability limits that states may increase but not reduce – provides a useful model for this tiered structure (Trout, 2024).

A further design consideration concerns the form of the capitalization requirement. Options include mandatory insurance, escrow accounts, surety bonds and parent-company guarantees (Montagnani et al., 2024). Each has advantages and disadvantages. Mandatory insurance is the most market-friendly approach but, as discussed in Section 5, faces significant actuarial challenges for AGI. Escrow accounts provide greater certainty of available funds but lock up capital that might otherwise be invested in safety research. A hybrid approach – requiring a base level of capitalization through escrow or surety bonds, supplemented by insurance for residual risk – may offer the best balance.

6.2 Dynamic duty-of-care standards: capability-calibrated obligations

The duty of care in negligence law is typically defined by reference to the standard of the reasonable person (or, in professional contexts, the reasonable professional) (Selbst, 2020; Wagner, 2019). For AI systems, the duty of care imposed on developers and deployers is defined by reference to regulatory obligations (under the AI Act) or general negligence principles. In either case, the standard is static: it is defined at a point in time and applied to the defendant's conduct at the time of the alleged breach (Soyer & Tettenborn, 2022).

AGI requires a dynamic duty-of-care standard that evolves in response to the system's capabilities. As an AGI system becomes more capable – through learning, self-modification or integration with additional data sources – the obligations of its developer and deployer should increase correspondingly (Faroldi, 2025). This principle could be operationalized through a system of capability thresholds, analogous to the risk tiers of the EU AI Act but calibrated to the specific challenge of general intelligence.

It is important to clarify how this capability-calibrated duty of care relates to the strict liability channeling proposed in Section 6.1 since the two are not in tension but operate at different levels (Dyson, 2022). Strict liability governs the operator's exposure to the victim: it determines who must pay compensation for AGI-caused harm, and removes fault, defect or foreseeability as a precondition of that claim. The dynamic duty of care, by contrast, governs two further questions that strict liability alone cannot resolve: first, the level of mandatory capitalization that the operator must maintain, which should scale with the system's measured capability tier; and second, the allocation of contribution or indemnity among the developer, deployer and any other actor in the AGI value chain once the operator has compensated the victim (Dyson, 2022). The operator's liability to the victim is therefore strict and capability-independent; the standard used to apportion ultimate responsibility within the value chain, and to calibrate the capitalization that backs that liability, is dynamic and capability-dependent.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

At the lowest tier, an AI system that demonstrates cross-domain capability in a limited number of domains would trigger enhanced disclosure obligations: the developer would be required to notify regulators and users of the system's expanding capabilities (Novelli et al., 2024). At intermediate tiers, cross-domain reasoning capability would trigger mandatory human-in-the-loop requirements for consequential decisions, enhanced post-market surveillance and periodic independent audits (Novelli et al., 2024). At the highest tier – systems demonstrating autonomous goal-setting and self-modification – the obligations would include mandatory containment protocols, real-time monitoring by independent auditors and pre-deployment safety certification by a specialized regulatory body (Novelli et al., 2024).

The legal basis for such dynamic standards exists in the EU framework: the AI Act's risk-based approach already calibrates obligations to the level of risk posed by the system. The innovation would be to make the calibration continuous rather than categorical, recognizing that AGI's capabilities are not fixed at deployment but evolve over time.

Implementing such a dynamic standard would require a measurement infrastructure for AI capabilities that does not currently exist in legally operationalizable form. The Morris et al. (2024) framework for "levels of AGI" provides a useful starting point, distinguishing between emerging, competent, expert, virtuoso and superhuman levels of AI performance. However, these levels need to be translated into legally actionable criteria: what specific benchmarks of cross-domain performance, autonomous reasoning or self-modification trigger which specific obligations? The development of such criteria should be a collaborative enterprise involving AI researchers, regulators and legal scholars, and the criteria should be updated regularly as the technology evolves.

The concept of dynamic duty-of-care standards also addresses a problem that the withdrawal of the AI Liability Directive has left unresolved: the relationship between regulatory compliance and civil liability (Bertolini, 2025b). Under the proposed Directive, compliance with the AI Act's requirements would have been relevant to assessing fault for high-risk AI systems, but would not have been conclusive. The dynamic duty-of-care approach proposed here goes further: it ties the standard of care directly to the system's measured capabilities, creating a clear and evolving benchmark against which developers' and deployers' conduct can be assessed. This approach provides greater legal certainty than the Directive's open-textured reference to a "duty of care" while retaining the flexibility to adapt to technological change.

6.3 International treaty-based compensation: learning from environmental liability

The international environmental liability regime offers a second instructive analogy. The International Convention on Civil Liability for Oil Pollution Damage (1969, as revised by the 1992 Protocol), the Fund Convention (1971, as revised by the 1992 Protocol) and the Supplementary Fund Protocol (2003) together establish a multilayered compensation system: primary liability is channeled to the ship-owner (strict liability with mandatory insurance), supplemented by an international fund financed by oil receivers, with an additional layer of supplementary compensation (Faure et al., 2015; Faure & Wang, 2008; Jacobsson, 1994). This structure ensures that victims can access adequate compensation even when the responsible party's resources are insufficient.

An international AGI compensation mechanism could be structured similarly (Cingillioglu & Lun, 2026). The first layer would comprise the AGI operator's mandatory capitalization reserves and insurance. The second layer would comprise an International AGI Compensation Fund, financed by mandatory contributions from entities that develop, deploy or profit from AGI systems. The third layer would comprise state-backed compensation, activated when the harm exceeds the capacity of the first two layers. This three-tiered structure addresses the jurisdictional fragmentation problem by creating a single international mechanism for compensation, regardless of where the harm occurs or where the AGI system was developed.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

The institutional home for such a mechanism would need to be determined through international negotiation. The existing international organizations with relevant competence include the Organization for Economic Co-operation and Development (OECD), which has already developed AI Principles (OECD, 2019, updated in 2024) and has institutional experience with technology governance; the United Nations, which established the High-Level Advisory Body on Artificial Intelligence in 2023 (United Nations, 2023); and potentially a new specialized agency modeled on the International Atomic Energy Agency, which combines international safety standards, safeguards/oversight functions and technical assistance. The Bletchley Declaration (2023) and the Seoul AI Summit Declaration (2024), the Paris AI Action Summit (2025), the UN Global Dialogue on AI Governance in July 2026 and the OECD's 2026 update to the Hiroshima AI Process voluntary reporting framework provide diplomatic precedents for international AI governance cooperation.

The treaty mechanism should also include provisions for technology assessment. One of the principal difficulties in regulating AGI is determining when a system has crossed the threshold from narrow to general intelligence (Faroldi, 2025). The proposed convention should establish an international technical committee – comprising AI researchers, ethicists and legal scholars – charged with developing and updating criteria for capability assessment. These criteria would determine when the enhanced obligations of the AGI liability regime are triggered. The committee's assessments should be binding on contracting states, subject to a review mechanism, to prevent regulatory arbitrage through divergent national interpretations of the threshold.

An important design choice concerns the relationship between the international compensation mechanism and existing national liability frameworks. The nuclear liability conventions adopt a principle of preemption: national courts apply the convention's rules rather than domestic tort law when the convention's conditions are met (McIntosh, 2022). For AGI, a similar preemption principle would have significant advantages, ensuring uniformity of treatment and preventing the forum-shopping that would otherwise result from the interaction of an international compensation mechanism with diverse national tort systems. However, preemption also carries risks, particularly if the convention's compensation limits are set too low or if its procedural mechanisms are less accessible than national courts. The design of the mechanism must balance these considerations, potentially through a minimum floor of compensation that states may supplement through national legislation.

6.4 The role of Singapore and ASEAN in an AGI liability architecture

The comparative analysis in Section 3 highlighted Singapore's distinctive institutional characteristics – its strong rule of law, sophisticated regulatory infrastructure and strategic position as a hub for technology companies operating across ASEAN – which give it a unique opportunity to play a leadership role in the development of AGI liability frameworks for the Asia-Pacific region. Singapore's existing governance frameworks – including the Model AI Governance Framework (PDPC, 2020), AI Verify (IMDA, 2022), the Model AI Governance Framework for Generative AI (IMDA, 2024, 2026a), the Model AI Governance Framework for Agentic AI (IMDA, 2026b) and related developments in national AI strategy, financial-sector AI risk management and AI testing and assurance (IMDA, 2026c; MAS, 2025, 2026; MDDI, 2026; Smart Nation Singapore, 2026) – provide a robust foundation upon which anticipatory liability provisions could be built. Singapore's internationally recognized emphasis on transparency, accountability and human-centric governance further strengthens that foundation (Cheong, 2024a).

Singapore has often served as a regulatory pioneer in the ASEAN context, developing governance frameworks capable of informing regional approaches. A similar dynamic could operate for

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

AGI liability: if Singapore were to develop a robust AGI liability framework – going beyond its current soft-law approach – it could establish a regional standard that facilitates both innovation and accountability.

The ASEAN Framework on Digital Data Governance (2018) and the ASEAN Guide on AI Governance and Ethics (ASEAN, 2024) provide a foundation for regional cooperation, with the latter reflecting Singapore's regional leadership through IMDA's role in its development. However, these instruments are nonbinding and do not address liability. An ASEAN-level agreement on AGI liability – perhaps modeled on the treaty mechanism proposed above but adapted to the region's developmental diversity – would complement the EU's approach and contribute to the emergence of a genuinely global governance architecture. Singapore, with its dual expertise in technology governance and international law, is well-positioned to lead this initiative.

7. Objections and responses

The proposals advanced in Section 6 will face predictable objections, which should be addressed candidly (Calo, 2015; Smuha, 2021).

7.1 The prematurity objection

The most common objection to AGI-specific regulation is that AGI does not yet exist, and it is premature to construct elaborate legal frameworks for a technology that may not arrive for decades, if ever (Rzeszucinski, 2024). This objection has force as a practical matter – regulatory resources are scarce and should be allocated to addressing present harms. Indeed, as Section 1 notes, most of the jurisdictions surveyed in Section 3 have yet to develop a fully coherent regulatory and liability framework even for ANI, the technology actually in deployment today; the case for diverting scarce regulatory attention toward a technology that does not yet exist is, on its face, correspondingly weaker (Faroldi, 2025). However, the history of technology regulation suggests that retroactive responses are both more costly and less effective than anticipatory ones. The environmental liability regime was largely constructed after decades of industrial pollution and major pollution incidents had caused widespread harm (Faure et al., 2015; Faure & Wang, 2008; Jacobsson, 1994). The nuclear liability regime, by contrast, was established at an early stage of commercial nuclear power and has proved more effective at ensuring compensation for victims than ordinary tort law alone (Schwartz, 2006).

Moreover, the anticipatory approach has informational benefits (Guston, 2014). The process of designing AGI liability rules forces a rigorous examination of the assumptions embedded in existing frameworks, generating insights that are valuable for governing current AI systems as well (Salmon et al., 2023). The five structural inadequacies identified in Section 4 are already present in embryonic form in current AI systems – the difficulty of defining defectiveness for learning algorithms (Wagner, 2019; Bertolini, 2025b), the challenge of establishing causation for opaque systems (Bathaee, 2018; Padovan et al., 2023) and the attribution problem in complex AI value chains (Beckers & Teubner, 2023; Kiener, 2025). Addressing these challenges at the AGI level illuminates their structure in ways that incremental reform of existing frameworks does not.

7.2 The innovation objection

A second objection, particularly prominent in the U.S. and Singapore contexts, is that anticipatory liability rules would stifle innovation by imposing excessive costs and uncertainties on AGI developers. This objection has some validity: the history of technology regulation illustrates the Collingridge dilemma (Collingridge, 1980), whereby early intervention risks being poorly informed while late intervention may be ineffective because the technology has already become entrenched. However, the argument cuts both ways. Inadequate liability rules also impede innovation by creating legal uncertainty, discouraging responsible development and concentrating development in jurisdictions with

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

the weakest protections (regulatory arbitrage) (Tartaro et al., 2023). The mandatory capitalization approach proposed above internalizes risk costs without prohibiting development, creating incentives for safety investment that are currently absent (Bertolini, 2025b; Erdélyi & Erdélyi, 2021; Verstappen, 2025).

The EU's experience with the AI Act is instructive. Although the EU's broader digital regulatory environment has been criticized for imposing excessive regulatory burdens and weakening competitiveness (Draghi, 2024), the Act reflects the EU's effort to combine innovation with a rights-based, trustworthy-AI framework, using regulation to steer technological development rather than merely constrain it (Smuha, 2021). An AGI liability framework that provides legal certainty and clear standards would similarly benefit developers by replacing the current uncertainty with predictable obligations.

7.3 The feasibility objection

A third objection concerns the feasibility of the proposed international treaty mechanism. International treaties are notoriously difficult to negotiate, slow to ratify and difficult to enforce. The existing international AI governance architecture – comprising nonbinding principles (OECD, 2019), summit declarations (Bletchley, Seoul and Paris) and voluntary commitments – reflects the practical challenges of achieving international consensus on technology governance (Maas, 2025).

This objection is well-taken but not dispositive. The nuclear and environmental liability conventions demonstrate that international treaty-based compensation mechanisms are feasible and effective, even in contexts involving divergent national interests and complex technical challenges (Bertolini, 2025b; Erdélyi & Erdélyi, 2021; Schwartz, 2006; Verstappen, 2025). The key is not to seek a comprehensive global agreement – which would be politically unrealistic – but to establish a coalition of willing jurisdictions (the EU, UK and like-minded states) that creates a critical mass of participation, with mechanisms for other jurisdictions to join over time. The mutual recognition provisions of the Paris Convention's nuclear liability regime provide a useful model (Tazaki, 2014).

The geopolitical dimension of AGI governance cannot be ignored. The development of frontier AI systems is concentrated in a small number of countries – primarily the United States and China, with significant activity in the UK, Canada and Israel and parts of Europe – and the competitive dynamics between these jurisdictions will likely shape the prospects for international cooperation. The U.S.–China rivalry in AI development creates particular challenges: both nations may have incentives to avoid liability regimes perceived as disadvantaging their domestic companies. However, this dynamic also creates opportunities: a coalition of middle powers – including the EU, UK, Japan, South Korea, Singapore and Australia – could help establish international norms that the major powers would face pressure to engage with, particularly if the coalition's market represents a significant share of global AGI demand (Calo & Lyu, 2026; Maslej et al., 2025).

7.4 The definitional objection

A fourth objection challenges the very possibility of defining AGI with sufficient precision for legal purposes. If AGI cannot be defined, the argument runs, it cannot be regulated, and any attempt to create AGI-specific liability rules will founder on definitional indeterminacy (Faroldi, 2025). This objection has particular force given the lack of consensus in the AI research community on the definition of AGI – or, indeed, on whether AGI is a coherent concept at all.

The response to this objection is to observe that legal categories need not correspond exactly to philosophical or scientific categories; they may be constructed instrumentally to serve regulatory, accountability and institutional purposes (Chesterman, 2021; Turner, 2018). Law routinely defines concepts for regulatory purposes in ways that depart from their usage in other domains. The EU AI Act's definition of an "AI system" (Article 3(1)) does not resolve the philosophical debate about the

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

nature of intelligence; it identifies a class of machine-based systems for the purpose of applying regulatory requirements (Smuha, 2021). Similarly, an AGI liability regime need not resolve the question of whether true general intelligence has been achieved; it need only identify specific capability thresholds – such as cross-domain reasoning, autonomous goal-setting or self-modification – that trigger specific legal obligations (Goertzel, 2014; Morris et al., 2024). The dynamic duty-of-care approach proposed in Section 6.2 is specifically designed to operationalize this capability-based approach, avoiding the need for a binary determination of whether a system is or is not AGI.

8. Conclusion

The existing global architecture of AI liability is the product of a specific historical moment: the emergence of narrow, task-specific AI systems that challenge existing legal frameworks in identifiable, manageable ways. The EU's revised Product Liability Directive, the now-withdrawn AI Liability Directive, the U.S.'s evolving tort jurisprudence and the soft-law approaches of jurisdictions like Singapore and the UK represent sophisticated responses to the challenges of narrow AI (discussed in Section 3). But they share a fundamental limitation: they are designed for systems that remain bounded by identifiable products, providers, deployers, users or sectoral contexts, and whose causal opacity is assumed to be manageable through disclosure, presumptions, regulatory guidance or ordinary evidentiary mechanisms.

AGI would strain each of these assumptions. Its domain generality destabilizes the concept of defectiveness. Its autonomous goal-setting weakens the link between human oversight and system behavior. Its causal opacity may exceed the capacity of any ordinary evidentiary mechanisms to reconstruct. Its self-modification dislocates the temporal framework of product liability (discussed in Section 2). Its inherently global operation would also expose the jurisdictional fragmentation that the withdrawal of the AI Liability Directive has already exacerbated within the EU.

This article has proposed an anticipatory liability architecture comprising three elements: mandatory capitalization reserves modeled on nuclear liability, dynamic duty-of-care standards calibrated to capability thresholds and an international treaty-based compensation mechanism modeled on environmental liability conventions (discussed in Section 6). These proposals are deliberately schematic; their elaboration would require extensive technical, economic and legal analysis beyond the scope of this article. But their fundamental premise is clear: the time to design liability frameworks for AGI is now, while the technology is still in development and the institutional and conceptual foundations can be established without the pressure of responding to catastrophic harm.

The history of liability law teaches that the most effective regimes are those that anticipate technological risk rather than respond to it. The nuclear liability conventions preceded the widespread deployment of commercial nuclear power. The environmental liability conventions were established in the wake of catastrophic oil spills but before the full scope of industrial environmental damage was understood. In both cases, the anticipatory approach enabled the development of institutional capacity, technical expertise and international cooperation that would have been far more difficult to establish after the fact (see Sections 6.1 and 6.3).

AGI presents an opportunity – and a responsibility – to apply this lesson. The alternative is to wait for an AGI system to cause catastrophic harm and then attempt to construct a liability regime amid political crisis, public outrage and the intense lobbying of entities with vast economic interests at stake. The legal academy, the policymaking community and international institutions should not wait for that moment. The work of building an anticipatory liability architecture for AGI must begin now.

Several practical steps can be taken immediately. First, the European Commission should issue guidance on how the revised Product Liability Directive should be interpreted in relation to highly autonomous AI systems, including AGI-specific risks, following the withdrawal of the AI Liability Directive. Even if comprehensive legislation is not politically feasible, guidance on the relationship

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

between product defectiveness, software updates, autonomous learning and highly autonomous AI systems would provide valuable legal certainty. Second, the OECD should commission a study on the application of nuclear and environmental liability principles to AGI, building on its existing work on AI Principles and the analytical framework developed in this article. Third, national law reform commissions – building on the UK Law Commission's pioneering 2025 paper – should systematically assess the adequacy of domestic tort law for highly autonomous AI and recommend specific legislative reforms.

Fourth, and perhaps most importantly, the AI safety and AI governance communities should deepen their engagement. The technical literature on AI alignment and the legal literature on AI liability have developed in relative isolation (Bertolini, 2025b; Buiten et al., (2023)). Yet the challenges they address are intimately related: alignment is the technical problem of ensuring that an AI system's goals are consistent with human values, and liability is the legal mechanism for allocating the costs when alignment fails. A liability regime that creates effective incentives for alignment investment – as the mandatory capitalization approach proposed above would do – serves both objectives simultaneously. The development of AGI liability frameworks should therefore be a collaborative enterprise, drawing on expertise from law, computer science, economics, philosophy and international relations (Perboli et al., 2025).

The transition from narrow to general AI is not a certainty, and its timeline remains deeply uncertain. But the history of technology governance teaches that legal frameworks are most effective when they are designed in advance of the technology they seek to regulate, not constructed in the aftermath of the harms they seek to prevent. The conceptual foundations of an anticipatory liability architecture for AGI can and should be laid now. This article represents a contribution to that project.

Acknowledgements. The author would like to express his sincere gratitude to Dr Monica Navarro-Michel and the two anonymous reviewers for their generous endorsement of this manuscript and their helpful suggestions, all of which have greatly improved the final version of this work.

Funding statement. No funding was received to assist with the preparation of this manuscript.

Competing interests. The author has no competing interests to declare that are relevant to the content of this article.

References

Allen, J. G., Loo, J., & Luna, J. (2025). Governing intelligence: Singapore's evolving AI governance framework. *Cambridge Forum on AI: Law and Governance*, *1*, e12. https://doi.org/10.1017/cfl.2024.12

American Law Institute. (1998). *Restatement of the Law Third*. Products Liability.

ASEAN. (2024). *ASEAN Guide on AI Governance and Ethics*.

Bathaee, Y. (2018). The artificial intelligence black box and the failure of intent and causation. *Harvard Journal of Law and Technology*, *31*(2), 889–938. https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-Yavar-Bathaee.pdf.

Beckers, A., & Teubner, G. (2022). *Three liability regimes for artificial intelligence: Algorithmic actants, hybrids, and crowds*. Hart Publishing.

Beckers, A., & Teubner, G. (2023). Human–algorithm hybrids as (quasi-)organizations? On the accountability of digital collective actors. *Journal of Law and Society*, *50*(1), 100–119. https://doi.org/10.1111/jols.12412

Bertolini, A. (2025a). *Intelligenza Artificiale e responsabilità civile: Problema, sistema, funzioni*. Il Mulino. https://www.amazon.it/Intelligenza-artificiale-responsabilit%C3%A0-Problema-funzioni/dp/8815391657

Bertolini, A. (2025b). Artificial Intelligence and Civil Liability: A European Perspective. European Parliament, Policy Department for Justice, Civil Liberties and Institutional Affairs, PE 776.426. July 2025. https://www.europarl.europa.eu/RegData/etudes/STUD/2025/776426/IUST_STU(2025)776426_EN.pdf

Bletchley Declaration. (2023). *The Bletchley Declaration by Countries Attending the AI Safety Summit*. 1–2 November 2023. Department for Science, Innovation and Technology.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. arXiv preprint arXiv:2303.12712.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

Buiten, M., de Streel, A., & Peitz, M. (2023).The law and economics of AI liability. *Computer Law & Security Review*, *48*105794. https://doi.org/10.1016/j.clsr.2023.105794.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, *81*, 77–91.

CAC. (2021). Provisions on the Management of Algorithmic Recommendations in Internet Information Services. 31 December 2021, effective 1 March 2022. Cyberspace Administration of China. https://www.chinalawtranslate.com/en/algorithms/

CAC. (2022). *Provisions on the administration of deep synthesis in internet information services*. 25 November 2022, effective 10 January 2023. Cyberspace Administration of China. https://www.chinalawtranslate.com/en/deep-synthesis/

CAC. (2023). *Interim measures for the management of generative artificial intelligence services*. 10 July 2023, effective 15 August 2023. Cyberspace Administration of China. https://www.chinalawtranslate.com/en/generative-ai-interim/

CAC. (2025). *Measures for labeling of AI-generated synthetic content*. 7 March 2025, effective 1 September 2025. Cyberspace Administration of China. https://www.chinalawtranslate.com/en/ai-labeling/

Cade Project. (2025). *China releases version 2.0 of AI security governance framework*. 16 September 2025. https://cadeproject.org/updates/china-releases-version-2-0-of-ai-security-governance-framework/

Calo, R. (2015). Robotics and the lessons of cyberlaw. *California Law Review*, *103*(3), 513–563. https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=1022&context=faculty-articles.

Calo, R., & Lyu, X. (2026). *Diverging AI governance structures in the Indo-Pacific*, 18 May 2026. National Bureau of Asian Research. https://www.nbr.org/publication/diverging-ai-governance-structures-in-the-indo-pacific/

Carvalho, A. (2026). *China's AI governance architecture: From state control to operational AI governance*. SSRN. https://doi.org/10.2139/ssrn.6656118

Cheong, B. C. (2022a). Avatars in the metaverse: Potential legal issues and remedies. *International Cybersecurity Law Review*, *3*(2), 467–494. https://doi.org/10.1365/s43439-022-00056-9

Cheong, B. C. (2022b). The historical origins of EU law primacy, its interaction with UK parliamentary sovereignty and Brexit consequences on other EU Member States. *European Papers*, *7*(3), 1447–1470. https://www.europeanpapers.eu/system/files/pdf_version/EP_eJ_2022_3_2_Articles_Ben_Chester_Cheong_00626.pdf.

Cheong, B. C. (2024a). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, *6*. Article 1421273. https://doi.org/10.3389/fhumd.2024.1421273

Cheong, B. C. (2024b). The rise of AI avatars: Legal personhood, rights and liabilities in an evolving metaverse. *Journal of Digital Technologies and Law*, *2*(4), 857–885. https://doi.org/10.21202/jdtl.2024.42

Chesterman, S. (2021). *We, the robots? Regulating artificial intelligence and the limits of the law*. Cambridge University Press.

Chopra, S., & White, L. F. (2011). *A legal theory for autonomous artificial agents*. University of Michigan Press.

Cingillioglu, I., & Lun, P. (2026). From AI to AGI: A human-in-the-loop framework for responsible IS futures. *Enterprise Information Systems*, Article 2653244. https://doi.org/10.1080/17517575.2026.2653244.

Collingridge, D. (1980). *The Social Control of Technology*. Frances Pinter.

Da Silva, M. (2024). Responsibility gaps. *Philosophy Compass*, *19*(9–10), e70002. https://doi.org/10.1111/phc3.70002

DeArman, A. (2019). The Wild, Wild West: A Case Study of Self-Driving Vehicle Testing in Arizona. *Arizona Law Review*, *61*, 983–1012. https://journals.librarypublishing.arizona.edu/arizlrev/article/id/6872/.

Department for Science, Innovation and Technology. (2025a). Data (Use and Access) Act 2025: Data protection and privacy changes. 27 June 2025. https://www.gov.uk/guidance/data-use-and-access-act-2025-data-protection-and-privacy-changes

Department for Science, Innovation and Technology. (2025b). AI Growth Zones. Collection, 13 November 2025. https://www.gov.uk/government/collections/ai-growth-zones

Department for Science, Innovation and Technology. (2025c). Delivering AI Growth Zones. CP 1440, 13 November 2025. https://www.gov.uk/government/publications/delivering-ai-growth-zones/delivering-ai-growth-zones

Department for Science, Innovation and Technology. (2026). Shaping Tomorrow: The UK's Digital Standards Strategy (2026 to 2030). 17 June 2026. https://www.gov.uk/government/publications/shaping-tomorrow-the-uks-digital-standards-strategy-2026-to-2030/shaping-tomorrow-the-uks-digital-standards-strategy-2026-to-2030

Digital Regulation Cooperation Forum. (2025). Call for Views: Agentic AI and Regulatory Challenges. 10 October 2025. https://www.drcf.org.uk/news-and-events/news/call-for-views-agentic-ai-and-regulatory-challenges

Dominguez, C., & McCauley, D. (2025). The international regime for nuclear liability: A key element of the regulatory framework for the transport of nuclear and radioactive material. International Atomic Energy Agency Conference Paper, ID, 193. https://conferences.iaea.org/event/421/papers/38042/files/14074-ID-193-Extended%20Synopsis%20The%20International%20Regime%20for%20Nuclear%20Liability%20.pdf

Draghi, M. (2024). *The future of European competitiveness (Part A)*. European Commission.

Duffourc, M. N. (2025). *The Withdrawal of the AI Liability Directive: A Critical Reflection on AI Liability in the EU*. Maastricht University Blog.

Duffourc, M. N., & Gerke, S. (2023). The proposed EU Directives for AI liability leave worrying gaps likely to impact medical AI. *Npj Digital Medicine*, *6*, 77. https://www.nature.com/articles/s41746-023-00823-w.

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

Dyson, M. (2022). Explaining the criminal and tortious developments in fault doctrine. In *Explaining Tort and Crime: Legal Development Across Laws and Legal Systems, 1850–2020* (pp. 220–296). Cambridge University Press. https://doi.org/10.1017/9781316534861.007

Ebers, M. (2021). Liability for artificial intelligence and EU consumer law. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law (JIPITEC)*, *12*(2), 204–220. https://www.jipitec.eu/jipitec/article/view/308.

Ebers, M., Janssen, A., & Meyer, O. (Eds.). (2009). *European perspectives on producers' liability*. De Gruyter.

Erdélyi, O. J., & Erdélyi, G. (2020). The AI liability puzzle and a fund-based work-around. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 50–56). https://doi.org/10.1145/3375627.3375806

Erdélyi, O. J., & Erdélyi, G. (2021). The AI liability puzzle and a fund-based work-around. *Journal of Artificial Intelligence Research*, *70*, 1309–1334. https://doi.org/10.1613/jair.1.12580

European Commission. (2022). Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence. COM(2022) 496 final.

European Commission. (2025). Communication from the Commission: A Competitiveness Compass for the EU. COM(2025) 30 final.

European Expert Group on Liability and New Technologies. (2019). *Liability for artificial intelligence and other emerging digital technologies*. Publications Office of the European Union.

European Parliament. (2017). Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

European Parliament. (2020). Civil liability regime for artificial intelligence (rapporteur: Axel Voss).

Faroldi, F. L. G. (2025). Risk and artificial general intelligence. *AI and Society*, *40*, 2541–2549. https://doi.org/10.1007/s00146-024-02004-z

Faure, M. G., Liu, J., & Wang, H. (2015). A multilayered approach to cover damage caused by offshore facilities. *Virginia Environmental Law Journal*, *33*, 356–422. https://cris.maastrichtuniversity.nl/ws/portalfiles/portal/72315351/faure_2015_a_multilayered_approach.pdf.

Faure, M. G., & Wang, H. (2008). Financial caps for oil pollution damage: A historical mistake? *Marine Policy*, *32*(4), 592–606. https://doi.org/10.1016/j.marpol.2007.10.008

Floridi, L., & Taddeo, M. (2018). Romans would have denied robots legal personhood. *Nature*, *557*, 309. https://www.nature.com/articles/d41586-018-05154-5.

Frattone, C. (2025). *Anatomy of a fall: On the anticipated withdrawal of the AI Liability Directive proposal*. Verfassungsblog. https://doi.org/10.59704/5f36c068798394a7.

Frazier, K. (2025). *We're not ready for AI liability*. AI Frontiers. https://ai-frontiers.org/articles/options-for-ai-liability

Giuffrida, I. (2019). Liability for AI decision-making: Some legal and ethical considerations. *Fordham Law Review*, *88*(2), 439–456. https://ir.lawnet.fordham.edu/flr/vol88/iss2/3/.

Gluck, J. (2025). Future of Privacy Forum, The State of State AI: Legislative Approaches to AI in 2025 (2 October 2025). https://fpf.org/blog/the-state-of-state-ai-legislative-approaches-to-ai-in-2025

Goertzel, B. (2014). Artificial general intelligence: Concept, state of the art, and future prospects. *Journal of Artificial General Intelligence*, *5*(1), 1–46. https://reference-global.com/article/10.2478/jagi-2014-0001.

Grace, K., Stewart, H., Sandkühler, J. F., Thomas, S., Weinstein-raun, B., Brauner, J., & Korzekwa, R. C. (2024). Thousands of AI authors on the future of AI. arXiv preprint arXiv:2401.02843.

Guston, D. H. (2014). Understanding 'anticipatory governance. *Social Studies of Science*, *44*(2), 218–242. https://journals.sagepub.com/doi/10.1177/0306312713508669.

Harris, L. A. (2025). Regulating Artificial Intelligence: U.S. and International Approaches and Considerations for Congress (Congressional Research Service Report No R48555, Library of Congress, 4 June 2025). https://www.congress.gov/crs-product/R48555

IMDA. (2022). *Singapore launches world's first AI testing framework and toolkit to promote transparency; Invites companies to pilot and contribute to international standards development*. Infocomm Media Development Authority of Singapore. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2022/sg-launches-worlds-first-ai-testing-framework-and-toolkit-to-promote-transparency

IMDA. (2024). *Singapore proposes framework to foster trusted Generative AI development*. Infocomm Media Development Authority of Singapore. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/public-consult-model-ai-governance-framework-genai

IMDA. (2026a). *Singapore launches new model AI governance framework for agentic AI*. 22 January 2026. Infocomm Media Development Authority of Singapore. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2026/new-model-ai-governance-framework-for-agentic-ai

IMDA. (2026b). *Updated model AI governance framework for agentic AI*. 20 May 2026. Infocomm Media Development Authority of Singapore. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/factsheets/2026/updated-model-ai-governance-framework-for-agentic-ai

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

IMDA. (2026c). *Singapore champions new global AI testing standardisation efforts on benchmarking and red teaming*. 20 April 2026. Infocomm Media Development Authority of Singapore. https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2026/singapore-champions-new-global-ai-testing-standardisation-efforts

Jacobsson, M. (1994). Compensation for oil pollution damage caused by oil spills from ships and the International Oil Pollution Compensation Fund. *Marine Pollution Bulletin*, *29*(6–12), 378–384. https://doi.org/10.1016/0025-326X(94)90657-2

Karvosky, H. (2025). Is State Farm Guilty of Racial Discrimination? The Presence of Algorithmic Bias in Claims Systems. Florida Undergraduate Law Review (3 July 2025). https://www.fulr.org/post/is-state-farm-guilty-of-racial-discrimination-the-presence-of-algorithmic-bias-in-claims-systems

Kiener, M. (2025). AI and responsibility: No gap, but abundance. *Journal of Applied Philosophy*, *42*(1), 357–374. https://doi.org/10.1111/japp.12765

Law Commission. (2025). *Artificial intelligence and the law: A discussion paper*. Law Commission of England and Wales.

Lemley, M. A., & Casey, B. (2019). Remedies for robots. *The University of Chicago Law Review*, *86*(5), 1311–1396. https://chicagounbound.uchicago.edu/uclrev/vol86/iss5/3/.

Library of Congress. (2023). China: Provisions on Deep Synthesis Technology Enter into Effect. 26 April 2023. https://www.loc.gov/item/global-legal-monitor/2023-04-25/china-provisions-on-deep-synthesis-technology-enter-into-effect/

Maas, M. M. (2025). The global AI governance architecture: Past and futures. In *Architectures of Global AI Governance: From Technological Change to Human Choice* (pp. 181–283). Oxford University Press. https://academic.oup.com/book/61416.

Madiega, T. (2023). European Parliamentary Research Service, Artificial Intelligence Liability Directive (Briefing, PE 739.342, February 2023) https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf

MAS. (2025). MAS guidelines for artificial intelligence (AI) risk management. 13 November 2025. Monetary Authority of Singapore. https://www.mas.gov.sg/news/media-releases/2025/mas-guidelines-for-artificial-intelligence-risk-management

MAS. (2026). MAS partners industry to develop AI risk management toolkit for the financial sector. 20 March 2026. Monetary Authority of Singapore. https://www.mas.gov.sg/news/media-releases/2026/mas-partners-industry-to-develop-ai-risk-management-toolkit-for-the-financial-sector

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., … Oak, S. (2025). The AI Index 2025 Annual Report. Stanford Institute for Human-Centered AI. https://hai.stanford.edu/ai-index/2025-ai-index-report

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, *6*, 175–183. https://link.springer.com/article/10.1007/s10676-004-3422-1.

McIntosh, S. (2022). Nuclear liability and post-Fukushima developments. In International Atomic Energy Agency (Ed.),*Nuclear Law* (pp. 249–269). T.M.C. Asser Press. https://link.springer.com/book/10.1007/978-94-6265-495-2

McLean, S., Read, G. J. M., Thompson, J., Baber, C., Stanton, N. A., & Salmon, P. M. (2023). The risks associated with Artificial General Intelligence: A systematic review. *Journal of Experimental and Theoretical Artificial Intelligence*, *35*(5), 649–663. https://doi.org/10.1080/0952813X.2021.1964003

MDDI. (2026). Update to Singapore's National AI Strategy: Refreshed priorities to harness AI for the public good. Factsheet, 20 May 2026. Ministry of Digital Development and Information. https://www.mddi.gov.sg/newsroom/update-to-singapore-s-national-ai-strategy–refreshed-priorities-to-harness-ai-for-the-public-good-factsheet/

Montagnani, M. L., Najjar, M.-C., & Davola, A. (2024). The EU regulatory approach(es) to AI liability, and its application to the financial services market. *Computer Law & Security Review*, *53*, 105984. https://doi.org/10.1016/j.clsr.2024.105984

Morris, A. (2025). Mobley v Workday: An Evolving AI Compliance Landscape. *Illinois Business Law Review*, *30*, 67–76. https://publish.illinois.edu/illinoisblj/files/2025/08/IBLJ-Summer-2025-V.30_Morris-1.pdf.

Morris, M. R., Sohl-dickstein, J., Fiedel, N., Warkentin, T., Dafoe, A., Faust, A., Farabet, C., & Legg, S. (2024). Position: Levels of AGI for operationalizing progress on the path to AGI. *Proceedings of the 41st International Conference on Machine Learning (PMLR 235* (pp. 36308–36321).

Mueller, M. (2025). AGI: The illusion that distorts and distracts digital governance. *Journal of Cyber Policy*, *10*(2), 140–157. https://doi.org/10.1080/23738871.2025.2597194

MultiState. (2026), State AI Legislation Tracker 2026: All 50 States. https://www.multistate.ai/artificial-intelligence-ai-legislation

Nannini, L. (2025). When Less Regulation Means More Complexity: The EU AI Liability Directive Withdrawal and Its Impact on European Technological Competitiveness. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, *8*(2), 1848–1861. https://ojs.aaai.org/index.php/AIES/article/view/36679.

Negele, M., Juijn, D., Shamir, A., Janku, D., Özcan, B., Soder, L., Velasco, L., Reddel, M., Bakker, M., Pacchiardi, L., & Andriushchenko, M. (2025). *Europe and the geopolitics of AGI: The need for a preparedness plan*. RAND Corporation. https://cfg.eu/wp-content/uploads/Europe-and-the-geopolitics-of-AGI.pdf

Novelli, C., Casolari, F., Hacker, P., Spedicato, G., & Floridi, L. (2024). Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity. *Computer Law & Security Review*, *55*, 106066. https://doi.org/10.1016/j.clsr.2024.106066

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

Novelli, C., Floridi, L., Sartor, G., & Teubner, G. (2025). AI as legal persons: Past, patterns, and prospects. *Journal of Law and Society*. https://doi.org/10.1111/jols.70021

NPC. (2020). Civil Code of the People's Republic of China. National People's Congress. Adopted 28 May 2020, effective 1 January 2021, arts 1202–1207 and 1236–1244. https://www.chinajusticeobserver.com/law/x/civil-code-of-china-part-vii-liability-for-tort-20200528

OECD. (2019, updated 2024). Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/449.

Oimann, A., & Tollon, F. (2025). Responsibility gaps and technology: Old wine in new bottles? *Journal of Applied Philosophy*, *42*(1), 337–356. https://doi.org/10.1111/japp.12763

Open Letter to the European Commission: Artificial Intelligence and Robotics. (2018). Artificial Intelligence and Robotics Experts. https://robotics-openletter.eu/

Padovan, P. H., Martins, C. M., & Reed, C. (2023). Black is the new Orange: How to determine AI liability. *Artificial Intelligence and Law*, *31*(1), 133–167. https://link.springer.com/article/10.1007/s10506-022-09308-9.

Papastergiou, F., Quintero, B., & Marin, V. (2026). Allocating responsibility in autonomous AI systems: A tiered governance model under EU regulation. *Social Sciences*, *15*(6), 392. https://doi.org/10.3390/socsci15060392

Paris AI Action Summit. (2025). *Paris AI action summit: Statement on inclusive and sustainable artificial intelligence for people and the planet*. Government of France.

PDPC. (2020). *Model AI Governance Framework* (2nd ed.). Personal Data Protection Commission.

Perboli, G., Simionato, N., & Pratali, S. (2025). Navigating the AI regulatory landscape: Balancing innovation, ethics, and global governance. *Economic and Political Studies*, *13*(4), 367–397. https://doi.org/10.1080/20954816.2025.2569584

Pirvan, P. (2025). *AI as product vs. AI as service: Unpacking the liability divide in EU safety legislation*. IAPP.

Rachum-twaig, O. (2020). Whose robot is it anyway?: Liability for artificial-intelligence-based robots. *University of Illinois Law Review*, *2020*(4), 1141–1176. https://illinoislawreview.org/wp-content/uploads/2020/08/Rachum-Twaig.pdf.

Reuters. (2024). China to develop more than 50 new standards for AI sector by 2026. 2 July 2024. https://www.reuters.com/technology/artificial-intelligence/china-develop-more-than-50-new-standards-ai-sector-by-2026-2024-07-02/

Rimkute, D. (2025). *AI liability after the AILD withdrawal: Why EU law still matters*. Oxford Business Law Blog, 1 April 2025.

Rodríguez de Las Heras Ballell, T. (2025). Mapping generative AI rules and liability scenarios in the AI Act, and in the proposed EU liability rules for AI liability. *Cambridge Forum on AI: Law and Governance*, *1*, e5. https://doi.org/10.1017/cfl.2024.8

Rzeszucinski, P. (2024). We will never achieve AGI. Towards AI, 20 February 2024. https://pub.towardsai.net/we-will-never-achieve-agi-c07033743ba6

Salmon, P. M., Baber, C., Burns, C., Carden, T., Cooke, N., Cummings, M., Hancock, P., McLean, S., Read, G. J., & Stanton, N. A. (2023). Managing the risks of artificial general intelligence: A human factors and ergonomics perspective. *Human Factors and Ergonomics in Manufacturing and Service Industries*, *33*(5), 366–378. https://doi.org/10.1002/hfm.20996

Santoni de Sio, F., & Mecacci, G. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy and Technology*, *34*, 1057–1084. https://link.springer.com/article/10.1007/s13347-021-00450-x.

Schwartz, J. A. (2006). International nuclear third party liability law: The response to Chernobyl. In OECD Nuclear Energy Agency & International Atomic Energy Agency, *International Nuclear Law in the Post-Chernobyl Period* (pp. 37–72). OECD. https://www.oecd-nea.org/upload/docs/application/pdf/2019-12/nea6146-iaea-chernobyl.pdf

Selbst, A. D. (2020). Negligence and AI's human users. *Boston University Law Review*, *100*, 1315–1376. https://www.bu.edu/bulawreview/files/2020/09/SELBST.pdf.

Selbst, A. D., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham Law Review*, *87*, 1085–1139. https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=5569&context=flr.

Seoul AI Summit Declaration. (2024). Seoul Declaration for Safe, Innovative and Inclusive AI: AI Seoul Summit, 21 May 2024. Government of the United Kingdom/Government of the Republic of Korea.

Sheehan, M. (2023). *China's AI regulations and how they get made*. Carnegie Endowment for International Peace.

Smakman, J. (2025). Risky business: an analysis of the current challenges and opportunities for AI liability in the UK (Discussion Paper, Ada Lovelace Institute, December 2025) https://www.adalovelaceinstitute.org/wp-content/uploads/pdfs/32827/risky-business.pdf?v=1770119947

Smart Nation Singapore. (2026). National AI Strategy. Updated 21 May 2026. https://www.smartnation.gov.sg/initiatives/national-ai-strategy/

Smuha, N. A. (2021). Beyond a human rights-based approach to AI governance: Promise, pitfalls, plea. *Philosophy and Technology*, *34*, 91–104. https://link.springer.com/article/10.1007/s13347-020-00403-w.

Soyer, B., & Tettenborn, A. (2022). Artificial intelligence and civil liability – Do we need a new regime? *International Journal of Law and Information Technology*, *30*(4), 385–397. https://doi.org/10.1093/ijlit/eaad001

Steel, S. (2024). Legal causation and AI. In E. Lim, & P. Morgan (Eds.), *The Cambridge Handbook of Private Law and Artificial Intelligence* (pp. 189–205). Cambridge University Press. https://doi.org/10.1017/9781108980197.010

Stempel, J., & Roy, A. (2025). Tesla ordered by Florida jury to pay $243 million in fatal Autopilot crash. Reuters, 1 August 2025. https://www.reuters.com/legal/litigation/tesla-ordered-by-florida-jury-pay-243-million-fatal-autopilot-crash-2025-08-01/

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, *42*(9), 1568–1580. https://www.sciencedirect.com/science/article/pii/S0048733313000930.

Tartaro, A., Smith, A. L., & Shaw, P. (2023). Assessing the impact of regulations and standards on innovation in the field of AI. arXiv:2302.04110. https://doi.org/10.48550/arXiv.2302.04110.

Tazaki, M. (2014). A nuclear third party liability regime of a multilateral nuclear approaches framework in the Asian region. *Sustainability*, *6*(1), 436–448. https://doi.org/10.3390/su6010436

The White House, Ensuring a National Policy Framework for Artificial Intelligence. (2025). https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/

Torrance, A. W., & Tomlinson, B. (2025). Decentral intelligence agency: The law and autonomous artificial intelligence. *Touro Law Review*, *40*(4). https://digitalcommons.tourolaw.edu/cgi/viewcontent.cgi?article=3519&context=lawreview

Trout, C. (2024). Liability and insurance for catastrophic losses: The nuclear power precedent and lessons for AI. arXiv:2409.06673v1. https://doi.org/10.48550/arXiv.2409.06673

Trout, C. (2026). When does regulation by insurance work? The case of frontier AI. arXiv:2512.06597. https://doi.org/10.48550/arXiv.2512.06597

Turner, J. (2018). *Robot rules: Regulating artificial intelligence*. Palgrave Macmillan.

UK Government. (2023). A pro-innovation approach to AI regulation. Cm 815. Department for Science, Innovation and Technology.

United Nations. (2023). High-Level Advisory Body on Artificial Intelligence. United Nations, New York.

US Senate Committee on the Judiciary. (2025a). Examining the Harm of AI Chatbots (Hearing, 16 September 2025). https://www.judiciary.senate.gov/committee-activity/hearings/examining-the-harm-of-ai-chatbots

US Senate Committee on the Judiciary. (2025b). Durbin, Hawley introduce bill allowing victims to sue AI companies (29 September 2025). https://www.judiciary.senate.gov/press/dem/releases/durbin-hawley-introduce-bill-allowing-victims-to-sue-ai-companies

Verstappen, J. (2025). AI Governance in Singapore. ESF Research Brief No. 3. Netherlands Enterprise Agency. https://www.rvo.nl/sites/default/files/2026-02/ESFResearchBrief3AIGovernanceinSingapore.pdf

Wagner, G. (2019). Robot, Inc.: Personhood for autonomous systems? *Fordham Law Review*, *88*, 591–612. https://ir.lawnet.fordham.edu/flr/vol88/iss2/8/.

Wagner, G. (2023). Liability rules for the digital age – Aiming for the Brussels Effect. Working Paper No. 10, Forschungsinstitut für Recht und digitale Transformation.

Weil, G. (2025). *The case for AI liability*. AI Frontiers. https://ai-frontiers.org/articles/case-for-ai-liability

Wendehorst, C. (2020). Strict liability for AI and other emerging technologies. *Journal of European Tort Law*, *11*(2), 150–180. https://www.degruyterbrill.com/document/doi/10.1515/jetl-2020-0140/html?srsltid=AfmBOopMuIeExrnSpOEcnvQe8-LNreNEcR_D0Lex6YJJbooD5So5hWj.

Yoshikawa, J. (2019). Sharing the costs of artificial intelligence: Universal no-fault social insurance for personal injuries. *Vanderbilt Journal of Entertainment and Technology Law*, *21*(4), 1155–1187. https://scholarship.law.vanderbilt.edu/jetlaw/vol21/iss4/8/.

Zhang, P. (2026). What do China's plans for 'comprehensive' new AI law mean for future of technology? South China Morning Post. https://www.scmp.com/news/china/politics/article/3353834/what-do-chinas-plans-comprehensive-new-ai-law-mean-future-technology

Dr Ben Chester Cheong, PhD & LLM (Cambridge), is a Lecturer of Law at the Singapore University of Social Sciences and a C-EENRG Fellow in the Department of Land Economy at the University of Cambridge. He has published extensively across company law, environmental law, legal theory and technology law in leading journals, including the Law Quarterly Review, Journal of Business Law, Journal of Environmental Law, Transnational Environmental Law, Capital Markets Law Journal and Statute Law Review. His work has been cited over 1,100 times on Google Scholar. An award-winning teacher, he received SUSS's Outstanding Teaching Award (2022) and Excellence in Teaching Award (2021).

Cite this article: Cheong, B.C. (2026). The liability vacuum: Why existing legal frameworks cannot govern artificial general intelligence and what must be done. *Cambridge Forum on AI: Law and Governance*, *2*, e19, 1–26. https://doi.org/10.1017/cfl.2026.10054

Downloaded from https://www.cambridge.org/core. 17 Nov 2026 at 10:20:56, subject to the Cambridge Core terms of use.