AI Defects in the New Product Liability Directive: A Conceptual and Economic Analysis

Working Chapter Draft.
Original Version: Jun 3, 2025; This version: August 27, 2025

Roee Sarel

Part I. Introduction

1 The European Union's new Product Liability Directive ('**PLD**'),[1] adopted in late 2024, is an intriguing attempt to adapt private law to the digital age. Motivated by fast technological change and increasing complexity of digital products,[2] the PLD introduces several significant changes to the harmonization of EU product liability law,[3] such as the expansion of eligible harm types,[4] presumptions that alleviate the burden of proof,[5] and differentiation between different 'economic operators' who may bear liability.[6]

2 This chapter deals with one change that stands out: the explicit application of the PLD to Artificial Intelligence (AI) products. The PLD mentions AI in several of its recitals[7] and adds the terms 'software' and 'digital manufacturing files' to the definition of a 'product.'[8] This change opens the door to a variety of questions surrounding defects in AI products.

3 The expansion of the PLD to AI products was originally intended to serve as part of a legislative triad, which includes (i) regulation of AI systems through a

[1] Directive (EU) 2024/2853 of 23 October 2024 on liability for defective products [2004], OJ L2024/2853 ('PLD').

[2] *S De Luca,* Revised Product Liability Directive, BRIEFING EU Legislation in Progress, <https://epthinktank.eu/2023/02/13/new-product-liability-directive-eu-legislation-in-progress/> accessed 30 March 2025 (referring to digital products' 'dependence on data [and] complexity and connectivity').

[3] *De Luca* (fn 2); *Z Jacquemin,* Product Liability Directive: Disclosure of Evidence, the Burden of Proof and Presumptions, Journal of European Tort Law (JETL) 2024, 126; *S Li/MG Faure*, The Revised Product Liability Directive: A Law and Economics Analysis, JETL 2024, 140.

[4] PLD, art 6(1)(c) and 6(2).

[5] PLD, art 10(2)-10(3).

[6] PLD, art 8.

[7] PLD, rec (3), (13), (40), (48).

[8] PLD, art 4(1) (defining product as including also 'electricity, digital manufacturing files, raw materials and software').

new AI Act,[9] (ii) the harmonization of product liability for AI products in the new PLD,[10] and (iii) a more general (fault-based) AI Liability Directive ('AILD').[11] However, while the first two pieces of legislation were adopted,[12] the proposed AILD was withdrawn by the European Commission in its 2025 program, citing 'no foreseeable agreement' as the grounds for withdrawal.[13]

4 These developments give rise to at least three interesting questions. First, what precisely does the PLD change for AI liability in the absence of its intended complement? Second, what does the current state of affairs—where liability for AI defects is subject to harmonization but other liability aspects are not—means for actors involved in the market for AI-based product? And third, does the PLD's liability regime for AI products yield efficient incentives for those actors.

5 The first question is mostly doctrinal and requires answering a series of sub-questions such as what constitutes an 'AI product' and when would such a product entail a 'defect.' In contrast, the second and third questions require a deep dive into the law and economics ('L&E') of AI liability, considering issues surrounding the optimal liability standard (eg strict liability or negligence), the optimal division of liability between the different economic operators along the AI's supply chain. This chapter strives to provide answers to these questions, building on a mixture of works on product liability, AI liability, the new PLD, and the L&E of AI. It partially relies on my own work in this area,[14] but extends the analysis to the specifics of the new PLD's final text and the question of dividing liability between multiple tortfeasors.

6 The analysis reveals that while the PLD explicitly brings AI systems within its scope, it creates significant interpretative challenges regarding when AI products are defective. In particular, the consumer expectation test used in the PLD proves particularly problematic for complex, autonomous, and self-learning AI systems where reasonable expectations are difficult to determine. More generally, from a L&E perspective, the PLD's liability regime creates some efficient and some inefficient incentives. It correctly implements strict

[9] Proposal for a Regulation Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM (2021) 206 final (Apr. 21, 2021) ('AIA Proposal').

[10] Proposal for a Directive of the European Parliament and of the Council on Liability for Defective Products, COM (2022) 495 final (Sept. 28, 2023) ('PLD Proposal').

[11] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), COM (2022) 496 final, (Sept. 28, 2022) ('AILD Proposal').

[12] Regulation (EU) 2024/1689 of the European Parliament and the Council of 13 June 2024 laying down harmonised rules on artificial intelligence, OJ 2024/1689 ('AI Act'); PLD.

[13] European Commission, Commission work programme 2025, COM(2025) 45 final;*S Li/M Faure*, Does the EU Need an Artificial Intelligence Liability Directive? Insights from the Economics of Federalism, Revue économique (RE) 2025, 115, 115.

[14] *R Sarel*, Restraining ChatGPT, UC Law Journal (UCLJ) 2023, 115.

liability with joint-and-several liability provisions for multiple tortfeasors, but its reliance on ambiguous tests, burden-shifting mechanisms, and defences like the development risk defence may lead to distorted innovation incentives. The absence of a complementary AI Liability Directive further complicates matters, potentially leading to inconsistent application across member states and strategic forum shopping. These challenges are particularly acute given the institutional limitations of courts in evaluating highly technical AI systems, creating a risk of either systematic over-deterrence or under-deterrence.

7 The rest of the chapter is structured as follows. Part II discusses the AI revolution and its relevance for product liability laws. Part III focuses on the definition of a defect and its application to AI systems. Part IV delves into who is liable for AI defects under the PLD. Part V entails a law & economics analysis of the current state of affairs, focusing on the optimal liability standard and optimal division of liability. Part VI concludes.

Part II. The AI revolution and its relevance to product liability

I. What is AI?

8 AI can be defined in many ways, but usually refers to some computational system capable of performing tasks that traditionally required human intelligence.[15] However, nowadays the term 'AI' often refers to 'probabilistic, large, resource-intensive machine-learning systems,'[16] and discussed in the context of Large Language Models (LLMs) and generative AI, including chatbots (eg OpenAI's ChatGPT, Google's Gemini, Microsoft's Co-pilot, Anthropic's Claude, and X's Grok) and content-generation apps that are operated using prompts.[17]

9 For the purposes of regulation, the EU adopted a specific definition of AI systems in its AI Act, namely: 'a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual

[15] See generally *GE Gignac/ET Szodorai,* Defining intelligence: Bridging the gap between human and artificial perspectives, Intelligence 2024, 101832.

[16] *DG Widder/M Whittaker/SM West,* Why 'open' AI systems are actually closed, and why this matters, Nature 2024, 827, 828.

[17] A 'prompt' is the input given by the user to the AI, which triggers the AI process that generates output. The art of designing prompt is known as 'prompt engineering'; see eg *MP Polak/D Morgan,* Extracting accurate materials data from research papers with conversational language models and prompt engineering, Nature Communications 2024, 1569.

environments.'[18] This definition emphasises two features: (i) autonomy and (ii) generated outputs based on inference from inputs.[19]

II. The benefits and risks of AI

10 AI offers substantial societal benefits through its capacity to process vast quantities of data, identify complex patterns, and automate routine processes with unprecedented efficiency.[20] Research demonstrates that AI applications are enhancing healthcare outcomes through improved diagnostic accuracy and personalized treatment protocols,[21] while simultaneously advancing scientific discovery by accelerating hypothesis testing and data analysis.[22] In economic contexts, AI systems can optimize resource allocation,[23] increase productivity,[24] and create new categories of employment opportunities even as they transform existing labour markets.[25]

11 Despite its many benefits, AI also entails substantial risks, which warrants careful consideration within legal and regulatory frameworks. For instance, AI may cause physical harm (eg an autonomous vehicle running over a pedestrian), economic harm (eg a faulty trading algorithm causing a client to lose money), emotional harm (eg an AI-algorithm nudging people into depression), or human rights infringements (eg privacy violations through unauthorized data processing or discrimination due to algorithmic profiling). Some of these harms are concentrated with specific victims, but others are widespread and highly dispersed. In particular, because AI eases the creation and spread of misinformation (fake news,

[18] AI Act, art 3(1).

[19] Note that autonomy is not synonymous with automation: 'an automated system functions independently but follows preprogrammed instructions, while an autonomous system possesses its own decisionmaking capacity' (*M Buiten,* Product liability for defective AI, European Journal of Law and Economics (EJLE) 2024, 239, 256).

[20] See eg *E Brynjolfsson/A Mcafee*, The business of artificial intelligence (2017) 7 Harvard Business Review 1; *Sarel*, UCLJ 2023, 115, 118; P *Kumar/D Choubey/OR Amosu/YR Ogunsuji/BE Abikoye/SC Umeorah*, Revolutionizing Sourcing with AI: Harnessing Technology for Unprecedented Efficiency and Savings, World Journal of Advanced Research and Reviews 2024, 1.

[21] *EJ Topol*, High-performance medicine: the convergence of human and artificial intelligence, Nature Medicine 2019, 44.

[22] H *Wang et al,* Scientific discovery in the age of artificial intelligence, Nature 2023, 47.

[23] *C Challoumis*, Building a sustainable economy-how ai can optimize resource allocation, in: XVI International Scientific Conference Proceedings (2024); *UF Ikwuanusi/C Azubuike/CS Odinu/AK Sule/U Francis*, Leveraging AI to address resource allocation challenges in academic and research libraries, IRE Journals 2022, 311.

[24] *R Seamans/M Raj,* AI, labor, productivity and the need for firm-level data (2018) NBER working paper no. w24239 <http://nber.org/papers/w24239> accessed 20 May 2025.

[25] *X Gao/H Feng*, AI-driven productivity gains: Artificial intelligence and firm productivity, Sustainability 2023, 8934.

deepfakes, etc), it may undermine public interests like social trust and the democratic processes.[26]

12 There is a vast body of literature that discusses the regulation and liability of AI, especially in light of the more recent technological leap in the field of generative AI.[27] The scholarly discourse encompasses a range of approaches, from risk-based regulatory frameworks to ethics-centered governance models, reflecting diverse perspectives on appropriate oversight mechanisms.[28] However, the discussion is often abstract and sometimes neglects the distinction between *general* liability and *product* liability. This distinction is important both conceptually and practically. Conceptually, one may wonder when is AI a 'product' and when can it be deemed 'defective,' especially when considering autonomous AIs that operate with no (or minimal) human involvement and are programmed to learn and adapt over time. And practically, the classification of AI as a product subjects it to product liability, which is characterized by a particular set of legal rules that do not always apply to general liability.

Part III. AI products and AI defects

A. When is AI a 'product'?

13 Notwithstanding the conceptual difficulties, the PLD is clear on its intention to expand the definition of a 'product' to all sorts of AI.[29] Article 4(1) includes "software" in the definition and recital (13) explains that software—including AI—should be considered a product irrespective of how it is supplied or used. It further clarifies that software may be either a standalone product or integrated as a component in another product, further supporting a wide definition.

14 However, recital (13) also states that *information* should not be considered a product, mentioning the 'content of digital files' and 'mere source code' as examples. It is not fully clear what this distinction actually means for AI. Consider a robot that relies on an algorithm, which makes use of a series of

[26] Cf *S Nasiri/A Hashemzadeh*, The evolution of disinformation from fake news propaganda to AI-driven narratives as deepfake, Journal of Cyberspace Studies 2024, 203, 203.

[27] Examples include: *Sarel*, UCLJ 2023, 115; *Buiten,* EJLE 2024, 239; *P Hacker,* The European AI liability directives–Critique of a half-hearted approach and lessons for the future, Computer Law & Security Review (CLSR) 2023, 1; *ME Kaminski*, Regulating the Risks of AI, Boston University Law Review 2023, 1347; *NG Packin/HY Jabotinsky,* Blocking as Regulating? Blacklisting Generative AI, American University Law Review 2024, 1467; M Herbosch, Liability for AI Agents, North Carolina Journal of Law & Technology (NCJOLT) 2025¸391.

[28] See generally *Sarel*, UCLJ 2023, 115*; HY Jabotinsky/R Sarel,* Co-authoring with an AI? Ethical dilemmas and artificial intelligence, Arizona State Law Journal 2024*,* 187.

[29] See *GI Grau*, The development risks defence in the digital age, European Journal of Risk Regulation (EJRR) 2025, 197, 198.

files. Are the robot and the algorithms both a 'product,' but the files are not? And if so, does product liability apply if the robot malfunctions because of a mistake in the files rather than in its algorithm?

15 In a ruling made under the old PLD's regime,[30] the Court of Justice of the European Union ruled that incorrect health advice appearing in a physical newspaper does not fall under product liability, because the information is effectively a service. However, in the new PLD, recital (17) explains that digital services are often integrated into (digital) products in a way that turns them into a feature of the product.[31] Presumably, this increases the PLD's scope to include also some types of information. For instance, outputs of LLMs, while technically "information" seem to be an inherent feature of AI-based products.

16 In an attempt to further address legal (un)certainty, recital (13) adds that a developer of an AI system *within its meaning of the AI ACT* should be treated as a manufacturer. Recall that the AI Act's definition of AI emphasises autonomy and the ability to generate outputs based on inference from inputs. This helps clarify some issues for the previously given example: the autonomous component of the robot will likely be classified as a product, but the data it relies on will not. This does not mean product liability is silent if such a robot causes harm, it simply means that claims of a defect should be aimed at the robot or its algorithm and not at the data. Of course, this does not mean that problems with data are never covered by the PLD, rather that the data itself is not a product.[32]

17 Next, recitals (39) and (40) try to tackle AI systems that adapt over time. These recitals jointly clarify that: significant modifications of a product should cause it to be treated as a new product, and that this holds even when the modification is 'due to the continuous learning of an AI system'.[33] This indeed eliminates

[30] CJEU 10.6.2021, C-65/20, *VI v KRONE – Verlag Gesellschaft mbH & Co KG*, ECLI:EU:C:2021:471. For further details, see *S Li,* The Definition of a Product: Between Software, Information and Services, in: *D Messner-Kreuzbaur* (ed.), The Revised Product Liability Directive. Open Questions at the Time of Implementation.

[31] See also PLD, recital (18), clarifying that related services should be considered under the manufacturer's control where they are "integrated into, or inter-connected with, a product…."; and PLD, recital (50), clarifying the control extends in cases that it takes the form of post-release updates or machine-learning algorithms.

[32] PLD, rec (30), says "Information is not…to be considered a product, and product liability rules should…not apply to the content of digital files, such as media files or e-books or the mere source code of software.". This hints at other forms of data that could be considered a product under the right circumstances (see also the wide definition of "data" adopted in art. 4(6), which refers to Regulation (EU) 2022/868 on European data governance of 30 May 2033, art. 2(1) ("'data' means any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording").

[33] PLD, rec. (40).

some additional uncertainty: AI falls under the PLD' scope even if it continuously updates. As AI would often fall within the PLD's scope, the more important question is would it be classified as 'defective.'

B. What is a defect?

18 The definition of a 'defect' varies across legal systems, but there are some commonalities. In the United States, defects are categorized into three types: (i) design defects, (ii) manufacturing defects, and (ii) warning defects (sometimes termed 'failure to warn,' 'failure to inform,' or 'informational defects').[34] Design defects concern the inherent conceptualization of a product rather than errors in its production.[35] Such defects occur when a product's blueprints create an unreasonable risk of harm, even when manufactured precisely as intended.[36] Manufacturing defects represent departures from the intended design specifications during the production process.[37] Warning defects involve inadequate communication about product risks.[38]

19 The difference between these types of defects is sometimes blurry. For example, is an inherently dangerous product defective by design, or does it become defective only when there is insufficient warning to the consumer? Yet there are two doctrinal tests that assist in determining whether a problem amounts to a 'defect' under US product liability law, one which focuses on consumers' expectations and another which ask whether there are reasonable alternatives with a reduced risk of harm (a 'risk-utility' test).[39] The US also provides some evidentiary shortcuts to alleviate the burden of proof in certain cases (eg when the incident that caused harm was of the 'kind that ordinarily occurs as a result of a product defect').[40] The main consequences of classifying a defect as belonging to a certain category is the standard of liability: manufacturing

[34] See eg *J Henderson/AD Twersky,* Achieving consensus on defective product design, Cornell Law Review (CLR) 1996, 867, 869; §2 Restatement (Third) of Torts: Prod. Liab. (Am. L. Inst. 1998) (defining a product as defective when 'at the time of sale or distribution, it contains a manufacturing defect, is defective in design, or is defective because of inadequate instructions or warnings.'). This restatement was introduced after an earlier restatement (Restatement of Torts, Second (Am. L. Inst., 1965) caused some confusion as to the differences between design defects and failure to inform (*J. Henderson/AD Twerski*, What Europe, Japan, and Other Countries Can Learn from the New American Restatement of Products Liability, Texas International Law Journal 1999, 34, 35.

[35] Cf. *DG Owen*, Design defects, Montana Law Review 2008, 215, 221.

[36] ibid.

[37] ibid.

[38] ibid 222.

[39] *CJ Masterman*/*WK Viscusi,* The specific consumer expectations test for product defects, Indiana Law Journal (ILJ) 2023, 183, 184-185.

[40] See Restatement (Third) of Torts: Prod. Liab. §3, which applies the *res ipsa loquitur* doctrine*,* where the plaintiff does not need not prove the defect in some cases; *Victor E. Schwartz*, The Restatement (Third) of Torts: Products Liability-The American Law Institute's Process of Democracy and Deliberation, Hofstra Law Review 1997, 743, 758.

defects are generally subject to strict liability, whereas design and warning defects are based on either strict liability or negligence, depending on the jurisdiction and type of lawsuit filed.[41]

20 Unlike the US, the EU does not categorize defects. Instead, the PLD simply dictates that 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.' Thereby, the EU sets two alternatives for establishing a defect: (i) a consumer-expectations text ('CET'),[42] and (ii) a legal requirement test. The CET was in force also in the old PLD, whereas the second test is a new addition.[43] However, it is not obvious whether the distinction between the two tests has practical relevance, as it seems hard to imagine a case where the law mandates a particular safety level (second test) that a person would not be entitled to expect (first test). Thus, for the remainder of this chapter, I will focus on the CET, which is relevant for both US and EU law and may anyway capture the special case of the legal requirement test.

21 The CET seems intuitive, but has long been debated, with some scholars arguing that it is too subjective and others claiming it is insufficiently subjective.[44] However, there are doctrinal filters that narrow down the scope of 'consumer expectations' as a benchmark, which may circumvent some of the issues. For example, the expectation is supposed to refer to 'reasonably expected use' and not just any use,[45] which introduces some objectivity into the evaluation.[46]

22 In addition, Article 7 of the PLD offers guidance for evaluating whether there is a defect, stating that 'all circumstances' should be considered and providing a list of particular circumstantial issues. This list references things such as the product's presentation and characteristics (labelling, design, etc), its reasonably

[41] See eg *AD Twersky*, Chasing the illusory pot of gold at the end of the rainbow: Negligence and strict liability in design defect litigation, Marquette Law Review 2006, 90 (explaining the difference between design defects based on negligence and those based on strict liability).

[42] A consumer expectation test was used also in the old PLD, see *Buiten,* EJLE 2024, 239, 252.

[43] Compare PLD, art 7 with Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products ('Old PLD'), art 6(1). See also PLD Proposal, art 6, which does not entail the second test.

[44] See eg *Masterman/Viscusi,* ILJ 2023, 183, 183 (proposing a 'specific consumer expectation test', which focuses on whether the defect increases the risk along the same dimension as the product's benefit); *Buiten,* EJLE 2024, 239, 254 (mentioning criticism that the consumer expectation test may be based on unreasonable expectations).

[45] *Buiten,* EJLE 2024, 239, 254.

[46] See also *Grau,* EJRR 2025, 197, 205 (arguing that the CET is evaluated based on a class of consumers and not a specific consumer and that this increases objectivity as well).

foreseeable use, and the specific needs of the group of users for whom the product is intended. Thus, one need not speculate generally what consumers (reasonably) expect but look into the specifics.

C. When is AI 'defective'?

23 Scholars have identified many challenges in determining whether an AI product is defected.[47] First, AI is often *complex,* which can make it difficult for courts to estimate what consumers positively expect or what consumers are entitled to expect as a normative manner.[48] For instance, do people expect autonomous vehicles to never have accidents, to be at least as good as human-driven vehicles, or something else?[49] Such questions are hard to answer also because consumers may not be experts on how AI operates or on how costly it is to achieve specific safety levels.[50] Consumers may also suffer from various cognitive biases that inhibit their ability to evaluate the risks that AI products pose.[51]

24 Second, AI systems are often *autonomous and self-learning,*[52] which means it can be hard to identify whether (i) something went wrong because of the AI's (autonomous) decisions, and whether (ii) human intervention would have provided better safety at a reasonable cost at any point in time. AI's autonomy also blurs the categorization of defects used in the US. For instance, generative AI is not only capable of drafting texts, but also programming new apps.[53] Suppose that a person 'A' uses an AI service 'B' to program another AI-based app 'C'. Assume that person A enters a prompt into service B, asking it to 'include all relevant safety warnings' when creating app C. Alas, the app C takes a decision that causes some harm. Would this be a design defect of app C (as person A erred when indirectly designing it using service B)? Would it be a manufacturing defect (as service B 'manufactured' the app C based on the design given by A's prompts)? Or would this be a warning defect (as A did not

[47] See *J De Bruyne/O Dheu/C Ducuing*, The European Commission's approach to extra-contractual liability and AI–An evaluation of the AI liability directive and the revised product liability directive, CLSR 2023, 1, 13.

[48] ibid; *Buiten,* EJLE 2024, 239, 255.

[49] Cf. *Hacker*, CLSR 2023, 1, 15 (discussing how AI can both make and avoid mistakes differently than humans and how this affects the evaluation of a defect in an autonomous vehicle).

[50] *Buiten,* EJLE 2024, 239, 259.

[51] See eg *Li/Faure,* JETL 2024, 140, 143.

[52] *Buiten,* EJLE 2024, 239, 255.

[53] Current examples include Cursor, Lovable, and Base44, but one can also use general AI chatbots like ChatGPT or Claude to create code.

make sure B's warnings in app C are sufficient)?[54] This toy example shows how autonomy complicates the analysis.

25 Yet, as AI can be involved in a wide range of products, it is important to keep in mind that not all AIs are equally problematic. Consider a voice-activated light bulb, in which AI is only used to decipher when a person in the room has said the words 'on' or 'off.' Such a device yields two groups of risks: those unaffected by the AI (eg the bulb's glass may break) and those affected by the AI (eg a failure of the AI to recognize the word 'off,' thereby keeping the light on for too long). For such simple cases, the risks seem predictable even if the AI itself is complex or autonomous. One way to address this point can be found in the PLD's recital (30),[55] which hints that high risks also yield higher safety expectations. Thus, in cases that are complex and entail substantial risks, the courts may determine that consumers' expectations were violated more easily, such that complexity and the degree of risk balance each other.

26 Third, AI raises particular difficulties in applying the CET. Some have gone so far as to argue in the context of AI that the CET is 'vague at best and bordering on a concept devoid of any content at worst.'[56] Apart from the aforementioned challenges of estimating what consumers (do or should) expect, there is also a need to figure out which product *uses and misuses* are foreseeable. To illustrate, consider the case of *Garcia v. Character Technologies,*[57] recently filed to a district court in Florida. The case concerns a 14 year old boy who committed suicide after engaging with an AI chatbot, allegedly after being nudged to do so by the AI. The plaintiffs are arguing that harms to minors from generative AI are generally foreseeable and that the defendant has failed in both mitigating the risks (a design defect) and providing proper warnings (a warning defect). In a motion to dismiss, the defendants raised several claims,[58] including that the chatbot is a service rather than a product and that a duty of care toward minors is owed only if the defendant has physical control over the minor. But suppose the defendants would have simply argued back that the AI is intended for adults,[59] such that the use by minors is a grave misuse that should exempt them

[54] Miriam Buiten argues that AI defects are mostly a consequence of design defects based on the idea that factors like 'training data, model architecture, learning algorithms, and decision-making rules' (*Buiten,* EJLE 2024, 239, 257). ,

[55] This point was identified by *Grau,* EJRR 2025, 197, 206, in his discussion of the PLD's draft before its adoption.

[56] *Hacker,* CLSR 2023, 1, 14.

[57] *M Garcia v Character Technologies Inc* [2025] US Dist Ct (MD Fla) Case No 6:24-cv-1903-ACC-UAM (20 May 2025) <https://www.courthousenews.com/wp-content/uploads/2025/05/garcia-v-character-technologies-order.pdf> accessed 23 May 2025.

[58] For instance, that liability for outputs of chatbots would violate the constitutional right of free speech.

[59] In the *Garcia* case, the app was marketed for ages 12+, so this claim does not help much.

from liability.[60] Such a claim would likely be rejected by the court, not only because use by minors seems foreseeable and thus covered by liability, but also because it is not unique to AI (ie all adult-oriented products may foreseeably by misused by minors under some conditions).

27 But what about other misuses? Suppose the defendants in *Garcia* would argue that (i) they could not foresee that their AI product would autonomously generate output that specifically encourages suicide and that (ii) this can only occur if the user deliberately manipulated the AI through highly unusual prompting techniques that no reasonable developer could anticipate. Such claims seem difficult to evaluate, but may well be very common in AI product liability litigation.

28 In a recent interim decision,[61] the court allowed the *Garcia* case to move forward and rejected the defendant's objections to the application of product liability to an AI chatbot. The court reasoned that lawsuits against such chatbots share the same logic as attacking design flaws, and that the plaintiff's claim of foreseeable risk when an AI chatbot is "released into the world" is sufficient for further examination.[62] The defendants' answer to the lawsuit indeed claims the harm is a result of "misuse, unauthori[s]ed use, unintended use, unforeseeable use and/or improper use" but without elaborating further, such that the court would likely have to determine how to evaluate forseeability.

29 In light of these difficulties, scholars have highlighted two potential solutions that can help sort out what constitutes a defect specifically in AI. The first solution involves replacing the CET with the aforementioned risk-utility test used in the US,[63] which considers alternative designs rather than expectations. Applying this test requires looking at factors such as the product's benefits, its safety features, safer alternative substitutes, the anticipated risks, the availability of warnings, and whether insurance can assist in spreading the loss.[64] Albeit this organized list of factors seems tempting, it is not obvious that it provides any more clarity than consumers' expectations. The reason is that the same issues that make AI so challenging would also interfere in figuring out these factors. For instance, how should courts evaluate which alternative designs would reduce the risk of an output encouraging suicide, such as in the

[60] *Character Technologies Inc*, Defendant Character Technologies, Inc.'s Answer and Affirmative Defenses to Plaintiff's First Amended Complaint in Garcia v. Character.AI (fn 57) <https://storage.courtlistener.com/recap/gov.uscourts.flmd.433581/gov.uscourts.flmd.433581.150.0.pdf> last accessed 27 June 2025.

[61] Garcia v. Character.AI (fn 57)

[62] The defendant's motion for Certification of Interlocutory Appeal (permission to appeal the interim decision) was rejected.

[63] *Buiten,* EJLE 2024, 239, 261; *Hacker,* CLSR 2023, 1, 14.

[64] *Buiten,* EJLE 2024, 239, 253, summarising the list by *JW Wade*, On the nature of strict tort liability for products, Mississippi Law Journal 1973, 827, 837.

Garcia case? Doing so requires inside information on the cost of development, the effectiveness of different algorithmic filters, or the ease in which one can (mis)use 'jailbreak' prompting to bypass the filters.[65] One might even need to go back to the data used to train the model and decide which items are most problematic.[66]

30 The second solution proposed by some scholars is based on the new PLD's list of circumstances that courts are asked to consider as part of 'all circumstances' in Article 7 (in fact, it has been argued that this list was created with AI in mind).[67] For instance, this list includes: 'the effect on the product of any ability to continue to learn or acquire new features after it is placed on the market or put into service,'[68] thereby acknowledging AI's attribute of self-learning. Some have deemed this list a 'meaningful step forward'[69] that will induce AI producers to pay attention to how their AI product develops over time. Others have raised concerns that this new addition would do just the opposite,[70] granting AI providers with an implicit defence of mitigating circumstances. The ambiguity as to whether the PLD's list of circumstances is meant to enlarge or shrink the scope of liability applies also to other items, leaving room for interpretation.[71]

31 Thus, albeit the PLD opens the door to a wide array of cases that might fall under the scope of AI defects, it does not provide fully clear answers as to how courts should deal with alleged AI defects.

D. Compensation (damages)

32 The PLD restricts the claimable damages to three categories: (i) death or personal injury (including medically recognised psychological harm); (ii) property damage (with a few exceptions), destruction or corruption of data used

[65] 'Jailbreaking' is the colloquial name for prompt engineering that tries to get LLMs to bypass their own safety requirements. See generally *Y Liu et al*, Jailbreaking chatgpt via prompt engineering: An empirical study (2023) arXiv preprint, <https://arxiv.org/abs/2305.13860> accessed 20 May 2025.

[66] For a recent proposal on figuring out whether an outcome of an LLM is attributable to a specific piece of data, see *N Vyas/S Kakade/B Barak*, On provable copyright protection for generative models (2023) International Conference on Machine Learning, 35277–35299 (proposing a so-called 'Near Access Free' framework that aims to elicit a counterfactual for the model without access to a specific item).

[67] *De Bruyne/Dheu/Ducuing*, CLSR 2023, 1, 13.

[68] PLD, art 7(2)(c).

[69] *Hacker,* CLSR 2023, 1, 14.

[70] *De Bruyne/Dheu/Ducuing,* CLSR 2023, 1, 13.

[71] *Buiten,* EJLE 2024, 239, 266.

for non-professional purposes.[72] The plaintiff can claim only material losses, unless national law permits to claim also non-material losses.[73]

33 This scope of damages deviates from the old PLD in two meaningful ways: First, the old PLD placed a minimum threshold of 500 EUR, which has been deleted in the new PLD. Second, harm to data is a new category added by the PLD, which is especially relevant for AI for obvious reasons.

E. The burden of proof

34 Under the PLD, the plaintiff generally needs to prove the defectiveness, the harm, and the causal link between the two.[74] However, the burden of proof is alleviated in several ways. For brevity, I will mostly focus on proving the defect (and mostly neglect the burden of proof surrounding the causal link).

35 The PLD enables the plaintiff to trigger a *rebuttable presumption* of a defect in four cases:[75] First, if the plaintiff has met a minimum evidentiary threshold of presenting 'facts and evidence sufficient to support the plausibility of the claim for compensation,' the plaintiff can ask the defendant to disclose relevant evidence at their disposal. If the defendant fails to do so, the plaintiff can ask the court to presume the defect on those grounds. Second, even if the defendant does comply with disclosure, the court can still presume the defect if either the plaintiff faces excessive difficulties (especially due to technical complexity) in proving the defect, or if the plaintiff has proven that the defect is 'likely.' Third, the plaintiff can demonstrates that the product does not comply with legal safety requirements that are also intended to protect the plaintiff from the relevant harm, which can trigger the presumption independently. Finally, the plaintiff can rely on a condition that resembles *res ipsa loquitur,*[76] by showing that the harm was caused by an obvious malfunction of the product during foreseeable use or ordinary circumstances. The defendant can try and rebut the presumption.

36 Having reviewed the main provisions of the PLD related to (AI) defects, the next part asks who is liable for the defect (that has been proven or presumed).

---

[72] PLD, art 6.

[73] See *Li/Faure,* JETL 2024, 140, 161 f.

[74] PLD, art 10(1).

[75] PLD, art 10.

[76] The *res ipsa loquitur* doctrine enables plaintiffs to establish negligence based on circumstantial evidence when the defendant had exclusive control over the instrumentality causing harm and the plaintiff did not contribute to their own injury; see generally *CE Carpenter,* The doctrine of res ipsa loquitur, University of Chicago Law Review 1934, 519. Similar doctrines exist in most, if not, all EU jurisdictions (see *C Kahn,* Product Liability Under Scientific Uncertainty : Does the New Directive Yield New Answers ?, in: *Messner-Kreuzbaur* (fn 3230)).

Part IV. Liability for AI defects under the new Product Liability Directive

37 Even if the plaintiff has proven that an AI is defective, it does not immediately follow that the defendant is liable. Rather, the PLD includes a series of general defences alongside rules surrounding the division of liability between different 'economic operators.' Let us review each in turn.

A. Liability of economic operators for AI defects

38 Article 8 of the PLD clarifies that the following economic operators are (potentially) liable:[77] the *manufacturer* of a defective product or component (with some limits),[78] the *importer* of a defective product or component, the *authorized representative* of the manufacturer, and—if there is no importer or authorized representative within the EU—the *fulfilment service provider*. In cases where seemingly none of the above exist, the plaintiff can sometimes ask to sue the *distributer* instead.

39 Each of these economic operators has a specific definition in the PLD, but let us focus on the more interesting case of a manufacturer. The directive first defines a manufacturer as someone who either (i) develops, manufactures or produces a product (for commercial purposes or their own use), or (ii) has the product designed or manufactured (including by putting their name or trademark on the product).[79] However, the PLD also adds that anyone who substantially *modifies* a product *outside of the manufacturer's control* and thereafter makes it available on the market is considered *as if* they are they are the manufacturer. The latter seems both useful and somewhat challenging for AI. It is useful because whenever someone significantly intervenes in the AI code (eg by releasing new updates). there is less ambiguity as to whether they are liable. And it is challenging because when changes occur through autonomous activity, it is not obvious whether they are ever truly outside the manufacturer's control.

40 The directive does include an explicit definition for 'manufacturer's control,' but it is not very helpful. Put briefly, a 'manufacturer's control' occurs if the manufacturer performs, authorises or consents to the integration, inter-connection or supply of a component (including software upgrades) or if it supplies the component directly.[80] While it clearly aims to be broad, it does not clarify the situation with respect to autonomous AIs, as the manufacturer does

---

[77] An 'economic operator' is defined as 'a manufacturer of a product or component, a provider of a related service, an authorised representative, an importer, a fulfilment service provider or a distributor' (PLD, art 4[15]).

[78] The liability for a defective component requires that it was 'Integrated into, or inter-connected with, a product within the manufacturer's control and caused that product to be defective.'

[79] PLD, art 4(10).

[80] PLD, art 4(5).

not directly consent ex-post to changes made by the AI, and it remains unclear whether the ex-ante delegation to the AI constitutes consent.

41 Otherwise, the PLD specifics that if two or more economic operators are liable for the same harm, they can be held liable *jointly and severally.*[81] For instance, if a defected AI product is designed by company A and then imported into the EU by company B, a plaintiff who suffers harm due to the defect can sue both A and B.[82] Moreover, member states are asked not to reduce the liability of an economic operator if the harm is caused by both a defect and an act or omission of a third party.[83] However, it does permit member states to adopt a contributory or comparative negligence regime, where liability is reduced or excluded if the injured person is at fault for their own harm.[84] Furthermore, if more than one economic operators are liable, then an operator who pays compensation has a right of recourse against the other liable operators (except a special case meant to protect SMEs).[85]

B. Liability exemptions

1. Various defences

42 The PLD provides specific defences in some cases where the plaintiff unjustifiably tries to drag multiple parties into the lawsuit. For example, manufacturers, importers and distributers are exempt if they show they did not actually put the product on the market.[86] A person who modifies a product is similarly exempted if they show the defectiveness is related to a part of the product unaffected by the modification.[87] An economic operator is also exempted if they show that the defect occurs due to compliance with legal requirements.[88] These defences are rather intuitive, but applying them to AI raises the same types of difficulties described above, as far as autonomous AIs are concerned (eg if an AI deploys itself following prompts from a user, is that user putting the AI on the market or not?). There are, however, several additional defences that seem more interesting for the discussion of AI defects.

---

[81] PLD, art 12(1).

[82] As an illustrative example, the interim decision in Garcia v. Character.AI (fn 57) found that Google, who supplied infrastructure for Character.AI could be viewed as a component manufacturer, but Google's parent company Alphabet could not.

[83] PLD, art 13(1).

[84] PLD, art 13(2).

[85] PLD, art 14, 12(2).

[86] PLD, art 11(a)-(b).

[87] PLD, art 11(g).

[88] PLD, art 11(d). Some have wondered whether compliance with regulatory standards provides a blanket defense, whereas non non-compliance trigger defectiveness almost automatically (also given the legal requirement test), for a discussion, see *G Wagner,* Next Generation EU Product Liability, JETL 2024, 172. 195 f.

2. The defect did not exist when the product was put on the market

43 First, the defendant can show it is probable that the defect did not exist at the time the product was put on the market. This defence seems especially tricky for self-learning AI, which changes over time, and it is not fully clear whether it actually adds anything. Recall that to identify a defect, one already has to consider the circumstances of learning over time and whether the moment was before or after the manufacturer had control. Thus, this defence seems only relevant if one first identifies a defect following the consideration of the timeline but then applies the defence following more or less the same reasons. From a practical perspective, it does not seem to matter much whether the defendant wins because they are exempted or because there is no defect at all, but there are possible exceptions (eg the symbolic role of identifying a defect may be important here).

44 Alas, the PLD makes things slightly more confusing on this front, as Article 11(2) adds an exception to the exemption, namely: an economic operator (who shows the defect probably did not exist when the product was put on the market) is again liable if the defectiveness is due to a related service, software or lack of software (including updates), or substantial modification of the product;[89] as long it is within the manufacturer's control. Admittedly, this exception is a bit confusing. Had the PLD fully excluded AI from the scope of defence, for instance, because it is too costly to keep track of when the defect was created, things would have been clear. Instead, one must locate precisely when the manufacturer's control ended, which then simultaneously determines (i) whether there is a defect at all, (ii) whether the defect is subject to a defence of (probably) not existing when the product was launched, and (iii) whether the exception to the exemption applies. This seems overly complicated, and perhaps having the manufacturer's control appear just once (explicitly or implicitly) would have been preferable.

3. Defect is attributable to the design

45 Second, a manufacturer of a defective component is exempted if the defectiveness is attributable to the design of the product in which the component was integrated or to instructions given by the product's manufacturer to the component's manufacturer.[90] This defence is intriguing, because it implicitly introduces a US-like distinction between design defects and manufacturing defects, where the latter grants an exemption in same cases that the former does not.

46 Yet for AI it is, once more, ambiguous. Suppose again that person A uses an AI service B and gives it a prompt to create an app C, which causes harm. Is the AI service B exempted because person A is the 'designer' who gave

[89] For a definition of 'substantial modification', see PLD, art 4(18).
[90] PLD, art 11(f).

instructions, or is service B the designer? Does it depend on the circumstances? Thus, this defence seems mostly helpful in simple cases (eg a robot designed by X and manufactured by Y) and less so in complex ones.

4. Development risk defence

47 A third relevant defence, known as the Development Risk Defence (DRD),[91] exempts a defendant from liability when the defect could not be discovered in real time given the 'objective state of scientific and technical knowledge at the time the product was placed on the market' *and* within the manufacturer's control. Importantly, one does not need to consider the manufacturer's subjective knowledge, only the *objective* knowledge.[92]

48 The DRD itself is not new and existed also in the old PLD (albeit the new PLD updates it by referring to the manufacturer's control). Historically, it emerged in response to a medical scandal surrounding the sale of a drug that had unintended negative consequences for pregnant women and their foetuses, and was controversial even at the time.[93] AI obviously presents different challenges than drugs, but its 'black box' nature, unpredictability, and ability to morph over time all seem highly relevant for the DRD.[94]

49 As AI producers cannot easily anticipate how autonomous AI will behave, some scholars have argued that the DRD should not apply to AI at all in order to avoid an implicit block exemption for all AI producers.[95] Other scholars have argued instead that digitization eases the access to scientific knowledge, so it will actually be very difficult for AI producers to successfully invoke the DRD.[96] In fact, it has been argued that when AI is deployed across many products and continuously learns from all units, it can become prohibitively difficult to invoke the DRD because simultaneous learning makes AI's behaviour predictable.[97] In my view, this argument is limited to very specific circumstances, where products have one clear use (eg autonomous vehicles are used for transportation) and the decision space is rather limited.[98] Otherwise,

[91] PLD, rec. (59); see also *Buiten,* EJLE 2024, 239, 252
[92] PLD, rec. (52).
[93] *Grau,* EJRR 2025, 197, 200 f.
[94] *Buiten,* EJLE 2024, 239, 266.
[95] See *Grau,* EJRR 2025, 197, 206; *P Machnikowski,* Producers' Liability in the EC Expert Group Report on Liability for AI, JETL 2020, 137, 146; *J-S. Borghetti*, Taking EU Product Liability Law Seriously: How Can the Product Liability Directive Effectively Contribute to Consumer Protection?, French Journal of Legal Policy 2023¸ 136, 179; *Kahn* (fn 76).
[96] *Grau,* EJRR 2025, 197, 204 f.
[97] Cf. *Grau,* EJRR 2025, 197, 207.
[98] Compare *D Messner-Kreuzbaur,* How Strict is EU Product Liability Now?, in: *Messner-Kreuzbaur* (fn 30) (proposing boundaries for the DRD, such as restricting it to knowledge about general natural facts).

continuous learning probably makes the AI's decision less predictable, as the parameters used for the decision keep changing.

50 While the new PLD's original draft did not include any additional instructions with respect to the DRD, the final version includes an additional article that permits member states to derogate from the DRD's scope. Article 18 allows member states to maintain existing provisions that exclude the DRD (ie an economic operator can be found liable even if the defect was undiscoverable given the state of scientific knowledge) and even adopt new provisions that do so, to a limited extent.[99]

51 Summing up, the new PLD sets out to explicitly address AI defects, adopting the CET as the relevant test and specifying circumstances for courts to consider. If reasonable consumers' expectation for safety are violated, court can hold multiple economic operators liable for the same harm, but these operators can try to invoke various defences. In the case of AI products, this setup yields many interpretative questions, especially for autonomous and self-learning AIs, such that it is far from clear when precisely courts would find economic operators of AI products liable. And since the AILD has been withdrawn, there is still much uncertainty.

52 With this in mind, the next part turns from the positive to the normative, asking not what the courts *will* do given the new PLD, but what the courts *should* do. The benchmark for evaluation will be mostly one of neoclassical L&E, assuming that the AI market is full of rational decision makers. However, I will also briefly consider some insights from behavioural L&E, that is, insights that assume deviation from perfect rationality.

Part V. The Law and economics perspective

I. Analytical steps: single tortfeasor

53 In previous work, I analysed how L&E can assist in determining how to best restrain generative AI,[100] focusing on a simple case with a single tortfeasor. That work addressed the EU's proposals for the PLD and the (now withdrawn) AILD, but the theory is general in nature. For the benefit of the reader, let me briefly summarize the relevant steps of the analysis (before proceeding to more complicated cases with multiple tortfeasors). Note that these steps are just one way of organizing the L&E discussion, i.e. there are not canonical.

---

[99] Member states adopting new provisions must only apply them to specific categories of products, based on justified public interest and in a proportional manner (PLD, art 18[22]). In addition, member states must notify the European Commission on such provisions and hold off on adopting the measure until the Commission consults with other member states and issues an opinion (within six months).

[100] *Sarel*, UCLJ 2023, 115.

A. Step 1: Market failures identification

54 Neoclassical economics assumes that people are selfish and perfectly rational —they are a 'homoeconomicus.'[101] In the absence of incentives to cooperate, people will, therefore, simply do what is best for them. Yet, if there are no frictions to trade (and in particular, in a perfect competitive market), this process leads to outcomes that are economically efficient. This outcome is the essence of Adam Smith's invisible hand: each person selfishly maximises their own utility and nonetheless this leads to a maximal social 'pie.'[102] With some straightforward analyses, this insight breeds two fundamental theorems (known as the first and second fundamental theorems of welfare economics): (i) competitive markets are efficient and (ii) if the government redistributes endowments (eg through taxes), people will trade with each other until they reach an efficient allocation again. The latter is expressed also by the seminal Coase Theorem, which uses a story of a farmer and a rancher who can trade at no cost (ie no frictions) and therefore reach an efficient deal.[103] The Coase theorem, therefore, argues that if transaction costs are sufficiently low, there is no need to intervene in the market.[104]

55 Because competitive markets are efficient, there is an economic justification for intervention only in cases where the market fails to deliver an efficient outcome. The usual suspects that cause markets to fail in that manner are (i) market power, (ii) asymmetric information, and (iii) externalities, where behavioural economics add a fourth one: (iv) behavioural failure, encompassing various cases with deviations from perfect rationality.[105]

56 Market power is problematic because sellers have an incentive to reduce the quantity produced and raise the price, which discourages purchases that could have yielded some utility for consumers. Asymmetric information is problematic because it may discourage sellers or buyers to go through with the transaction, fearing that they are dealing with a counterparty whose product is of too low quality (an 'adverse selection problem') or who will breach the contract because monitoring their compliance is too difficult (a 'moral hazard

---

[101] *J-P Elm/R Sarel,* No Policy is an Island: Mitigating COVID-19 in View of Interaction Effects, American Journal of Law & Medicine (AJLM) 2022, 7, 22.

[102] ibid 14.

[103] *Cooter/Ulen* (fn. 118) 81; *R Sarel,* Property rights in Cryptocurrencies: A Law and Economics Perspective, NCJOLT 2021, 389, 422-423. See also *B Köksal/R Sarel*, The Smart Contracts Trilemma, University of Illinois Law Review 2025 (forthcoming), 101, 139-140. The theorem is based on Coase's classic paper (*RH Coase*, The Problem of Social Cost, Journal of Law & Economics 1960, 1), but there is debate on its content (*SG Medema*, A case of mistaken identity: George Stigler, "The Problem of Social Cost," and the Coase theorem, EJLE 2011, 11).

[104] See eg *Li/Faure,* JETL 2024, 140, 142 f.

[105] *HY Jabotinsky/R Sarel*, How crisis affects crypto: Coronavirus as a test case, Hastings Law Journal (HLJ) 2023, 433, 452.

problem').[106] Externalities are effects on third parties, which cause the parties to either trade too much (if there is a negative externality) or too little (if there is a positive externality), given that they simply do not care about third parties' welfare.[107] Behavioural market failures include, for instance, cases of 'herding', where consumers mimic each other's trading strategies due to cognitive biases and lead to bad outcomes like price bubbles.[108]

B. Step 2: Choosing a legal tool

57 If a market failure occurs, the government can decide to address it using different interventions, including public law (eg regulation like the AI act), private law (eg contract law, tort law, or unjust enrichment law), and some other options (eg taxes). Tort law, and product liability law in particular, is mostly relevant for the case of negative externalities.

58 However, not all negative externalities should be dealt with using tort law. For example, a seminal paper by Steven Shavell argues that defendants who anticipate they will easily win lawsuits (or not be sued at all) will not respond to the threat of tort liability.[109] Such defendants include those who are 'judgment proof' because they are insolvent but also those who cause harms that (i) are widely dispersed, (ii) manifest only in the future, or (iii) involve an ambiguous causal link. Shavell argues that such cases are, ceteris paribus, better dealt with regulation than liability.

59 It is easy to see how AI products may cause negative externalities. Economic operators whose primary goal is profit maximization will not directly care about the welfare of third parties (or even of their own consumers),[110] so in the absence of liability, they would have no direct incentive to implement safety features. Negotiation with third parties may be infeasible because the operators cannot easily anticipate who the victims of their AI products will be (eg who will possibly be hit by an autonomous vehicle, whose data will be lost, etc) or how high the harm will be. Given negative externalities and high transaction costs, L&E generally recommends using liability rules.[111]

---

[106] *Cooter/Ulen* (fn. 118) 48.

[107] *Cooter/Ulen* (fn. 118) 39.

[108] *HY Jabotinsky/R Sarel*, HLJ 2023, 433, 447.

[109] *S Shavell,* Liability for harm versus regulation of safety, The Journal of Legal Studies (JLS) 1984, 357.

[110] Economic operators might care about reduced profitability if their neglect of safety features would reduce the demand of consumers, but not all users are consumers and it is not obvious that the decrease in demand would be sufficient to matter here.

[111] See generally the seminal paper by *G Calabresi/AD Melamed,* Property rules, liability rules, and inalienability: one view of the cathedral, Harvard Law Review 1971, 1089. If the negative externalities are extremely high, the recommendation might be instead to use inalienability rules that simply prohibit the transaction (ibid).

C. Step 3: Choosing a liability regime

60 If liability rules are recommended, one must make some choices on how to design them. Traditionally, L&E divides the discussion into three groups of costs: primary costs (the cost of the accident), secondary costs (who bears the risk), and tertiary costs (administrative costs).[112] For brevity, I will summarize only the main points on how legal rules can address those costs.

61 The most heavily discussed issue is the choice of liability standard, mostly comparing fault-based (eg negligence) and non-fault-based (eg strict liability) standards. In the context of (generative) AI and a single tortefeasor, I have previously discussed the key considerations: (i) unilateral versus bilateral care, (ii) activity levels, (iii) insurance, and (iv) known versus unknown risks.[113]

62 In a nutshell, in unilateral care cases (where only the tortfeasor can affect the likelihood of harm), negligence and strict liability can lead to the same—and efficient—level of care. For strict liability, this is straightforward: the tortfeasor simply always pays for the harm, which causes them to fully internalize it. Thus, under strict liability, the single tortfeasor behaves as if they maximize social welfare. The reason why the same outcome emerges under negligence is slightly different and relies on the famous 'Learned Hand Formula.'[114] This formula suggests that a tortfeasor should be held liable whenever their failure to take precautions is inefficient, that is, when the precaution's costs (usually denoted 'B') is lower than the expected harm (usually denoted L*P, where L is the size of the harm and P is the probability of the harm occurring). Hence, by definition, a tortfeasor is deemed negligent if they did not maximize social welfare, yielding the same outcome as strict liability. Interestingly, the regime of no liability, which typically yields under-deterrence, sometimes leads to the same result as negligence and liability in the case of product liability but only under two conditions:[115] (i) if the victims are *all* consumers and (ii) if the consumers-victims are perfectly informed about the harm. In this case, consumers will incorporate their harm into their willingness to pay. Consequently, a tortfeasor-seller will prefer to pay the cost of precautions whenever they are lower than the expected harm, as this will increase their profits through higher prices.[116] But in any other case (third party victims, imperfectly informed consumers), a no liability regime yields under-deterrence.

63 For bilateral care cases, the conclusion differs because the victim can also influence the likelihood of harm. In that case, the standard of liability is only efficient if it incentivises also the victim to take efficient care. This occurs if the

---

[112] *G Calabresi*, The Costs of Accidents: A Legal and Economic Analysis (1970).
[113] *Sarel*, UCLJ 2023, 115, 134 ff.
[114] *Cooter/Ulen* (fn. 118) 214.
[115] See *Buiten,* EJLE 2024, 239, 244.
[116] *S Shavell*, Foundations of economic analysis of law (2004) 213.

tortfeasor is subject to negligence but not strict liability. The reason is that under negligence, the tortfeasor is exempted from harm once they take efficient precautions, leaving the victim to bear the remaining risk. The victim then has all the incentives to minimize that remaining risk by taking efficient precautions. In contrast, strict liability grants the victim with an implicit insurance: the tortfeasor always pays for the harm, the so victim has no incentive to reduce it. [117]

64 Activity levels are a quantification of the tortfeasor and the victim's intensity of actions. For example, a tortfeasor can produce one product (low activity) or many products (high activity) whereas a victim can use the product once or frequently. Negligence rules generally do not capture activity levels and therefore may induce inefficient actions, as they exempt from liability anyone who takes precautions, irrespective of how much risk the person causes overall. On the contrary, strict liability induces the tortfeasor to internalize everything, including activity levels. But on the flip side, strict liability grants the victim with the aforementioned implicit insurance, which gives 'bad' incentives also for the victim's activity levels. Some of these issues can be addresses through more creative liability standards, such as those that add contributory or comparative negligence or those that use proportional liability, but there is no 'silver bullet' that solves all the problems simultaneously.[118]

65 Importantly, the efficiency of a liability regime (whether strict liability or negligence) also depends on an underlying assumption: that the legal system can accurately distinguish between socially harmful and socially beneficial conduct. If the courts systematically misclassify efficient behaviour as defective—reducing the payoff differential between good and bad conduct—it can undermine the very incentive structure that makes liability regimes efficient in the first place.[119]

66 Otherwise, L&E also highlights the different effects of market insurance, such as (i) the concern that an insured party will not take precaution because they are insured, and (ii) the need to incentivize the parties to buy market insurance, when one is available and it is efficient to do so. A parallel discussion concerns the need to incentivize parties to seek out information that can assist in preventing the harm whenever the risks are unknown.

---

[117] *Sarel*, UCLJ 2023, 115, 142 f.

[118] *R Cooter/T Ulen*, Law and Economics (6th edition, 2016) 204 (see table 6.2).

[119] A helpful analogy can be drawn to work on wrongful convictions, which shows how punishing benign behavior (here, classifying non-defects as defects) incentivizes a switch to the alternative—harmful behavior (see *R Sarel,* Crime and Punishment in Times of Pandemics, EJLE 2022, 155, 160 ff).

67 More generally, these different discussions reflect the concept of the *Least Cost Avoider* (LCA)[120]—ensuring that the party who can prevent the harm for the lowest cost is incentivized to do so—and the *Least Cost Information Gatherer* (LCIG)[121]—ensuring that the party who can attain the relevant information for the lowest cost is incentivized to do so. Sometimes these go together (eg a large polluting firm who can both prevent the pollution and has easy access to experts) and sometimes they are at odds (eg when the victim is the expert).

D. Step 4: Procedural and institutional choices

68 The analysis of optimal liability should yield a general candidate for a liability rule (eg negligence or strict liability), but there are other considerations that matter. These include issues such as (i) strategic substitutes or complements (are there other rules in place that assist or get in the way of the liability rules? Do public enforcers have an incentive to increase or reduce their effort if private enforcement is in play? etc), (ii) the attributes of the courts (do judges have an incentive to apply the liability standard correctly? Do courts have access to the relevant experts? etc), and (iii) the burden of proof (how easy it is to prove fault? When does the burden shift to the defendant? etc).

69 The existence of strategic substitutes (eg other legal policies serving the same purpose) means the problem may already be addressed, such that the benefits of liability is lower. In contrast, the existence of strategic complements (other policies that become more effective in conjunction with liability) means the benefits could be higher.[122]

70 Application by judges is also important when moving from theory to practice, for instance, because judges may care about more than just applying the legal standard per se, for instance, because they want to save on effort costs, avoid being reversed on appeal, or implement their preferred ideology.[123]

71 The burden of proof further matters for incentives because requiring the plaintiff to prove a full causal link may cause under-deterrence (as the defendant anticipates winning the lawsuit due to the difficulty to prove the link) but presuming the causal link may cause over-deterrence (as the defendant may

---

[120] See eg *Sarel*, UCLJ 2023, 115, 148; *G Dari-Mattiacci/N Garoupa*, Least-cost avoidance: the tragedy of common safety, The Journal of Law, Economics, & Organization 2009, 235; sometimes the term 'cheapest cost avoider' is used instead (eg *E Carbonara/A Guerra/F Parisi*, Sharing residual liability: the cheapest cost avoider revisited, JLS 2016, 173).

[121] *Sarel*, UCLJ 2023, 115, 148; A-S Vandenberghe, The role of information deficiencies in contract enforcement, Erasmus Law Review 2010, 71, 76.

[122] See generally Elm/Sarel, AJLM 2022, 7, 22.

[123] See generally *E Feess/R Sarel,* Judicial effort and the appeals system: Theory and experiment, JLS 2018, 269; *R Sarel/M Demirtas*, Delegation in a multi-tier court system: are remands in the US federal courts driven by moral hazard?, European Journal of Political Economy 2021, 1.

chill their activity to avoid distant harms that are not a direct result of their actions).[124]

II. Analytical steps: multiple tortfeasors

A. Join torts

72 If the case of multiple tortfeasors, L&E offers more intricate distinctions. For instance, there are differences between a 'simultaneous' joint tort, where multiple parties cause a single (or indivisible) injury at the same time, and a 'successive' joint tort, where one tortfeasor causes harm and another tortfeasor aggravates it.[125] Within simultaneous joint torts, one can further distinguish between 'alternative care,' where each tortfeasor can unilaterally prevent the entire harm, and 'joint care,' where care by multiple tortfeasors is required to prevent the harm. There are also differences between joint-and-several liability, where each tortfeasor must pay for the share of other tortfeasors who are insolvent, and non-joint liability, where each tortfeasor's liability is limited to their share.[126] Furthermore, it may matter whether tortfeasors are subject to a 'contribution' rule, where they have to indemnify one another (under joint-and-several liability) according to some division rule, or a 'no contribution' rule, which excludes indemnification.[127]

73 Yet there is an ongoing debate on whether all such distinctions matter. As a starting point, consider an old economic argument saying that optimal deterrence requires holding every tortfeasor liable for the entire harm they cause, irrespective of whether other tortfeasors are involved.[128] To see why this argument might make sense, consider the following simple variant of the Hand formula: Assume there are multiple (but homogeneous) tortfeasors who can efficiently prevent the harm, but that each tortfeasor only pays for a fraction $\alpha \in [0,1]$ in damages to the victim. A rational tortfeasor who maximises their utility will only invest in precautions if the cost of doing so (B) is lower than the expected payment of damages, which now can be denoted as $\alpha * LP$ (where LP is the expected harm). The tortfeasor will only invest if $B < \alpha LP$ but the Hand formula requires they invest if $B < LP$. This creates a mismatch, which only converges if $\alpha = 1$, that is, if the tortfeasor pays full damages. Yet there is a cost to doing so: if all tortfeasors are induced to invest in precautions, this

[124] *Sarel*, UCLJ 2023, 115, 171 f.

[125] *WM Landes/RA Posner*, Joint and multiple tortfeasors: An economic analysis, JLS 1980¸517, 517.

[126] *J Jacob/B Lovat*, Economic analysis of liability apportionment among multiple tortfeasors: A survey, and perspectives in large-scale risks management, Chicago-Kent Law Review (CKLR) 2016¸659, 661.

[127] See generally *Landes/Posner,* JLS 1980¸ 517; *LA Kornhauser/RL Revesz,* Sharing damages among multiple tortfeasors, Yale Law Journal (YLJ) 1989, 831.

[128] This argument dates back to *AC Pigou*, The Economics of Welfare (4th edn 1932); see *Jacob/Lovat*, CKLR 2016¸659, 659 f.

may lead to wasteful duplicative investments. But now suppose instead that tortfeasors have heterogeneous precaution costs, such that some can prevent it very cheaply while for others prevention is expensive. In such a case, inducing investment from all tortfeasors is not only wasteful but also inconsistent with the LCA principle.[129] Thus, when there exists a least cost avoider, they should bear full liability either ex-ante (by exempting everyone else) or ex-post (by allowing other tortfeasors to seek our indemnification from the least cost avoider).[130]

74 But most of the debate stems from differences in assumptions. For instance, Landes and Posner consider a setting where there are two tortfeasors who (i) face a negligence standard, (ii) cannot seek out indemnification from each other, and (iii) are liable for different fractions of the harm that sum up to 1 (say tortfeasor 1's share is α_1 and tortfeasor's share is $\alpha_2 = 1 - \alpha_1$). If it is efficient for both tortfeasors to prevent the harm ($B_1, B_2 < LP$), then at least one of them has an incentive to invest in precautions and thereby to avoid liability. Anticipating that, the other tortfeasor will also invest.[131] They give the following numerical example:[132]

> 'Suppose that A and B share a party wall and that C will be injured if it collapses because A and B fail to maintain it adequately. The expected accident cost to C is $100 and the optimal expenditure on maintenance by A and by B is $40 each. The legal rule is no [indemnification], and let us assume that if A and B spend nothing on maintenance the expected liability of each is $50, i.e., one-half the expected damages of C.' A and B then know that by spending $40 each can reduce his expected liability to zero, and knowing that each will spend the $40. The assumption that optimality requires equal expenditures on accident avoidance by the potential injurers is inessential. Suppose the optimal method of avoiding the $100 expected accident cost is for A to spend $79 and B $1, and as before the expected liability cost of A and B is $50 each. B will spend $1 to avoid an expected liability cost of $50; and once he has done so, A will be faced with the prospect of having to pay $100 unless he spends $79 on care. Since $79 is less than $100, he will be careful too. The sequence will fail only if the sum of A's and B's avoidance costs exceeds $100 (regardless of the optimal division of those costs between A and B), but in that event they would not be negligent in failing to take the precautions'

75 Landes and Posner further argue that a no-contribution rule and a contribution rule both lead to efficient outcomes, but the latter has the disadvantage of higher

---

[129] Cf. *M Buiten/A de Streel/Martin Peitz,* The Law and Economics of AI Liability, CLSR 2023, 1, 11.

[130] *Landes/Posner,* JLS 1980¸ 517, 526.

[131] Landes and Posner eventually reach the same conclusion for both simultaneous and successive joint care, as long as the latter holds the first tortfeasor liable for the entire harm and the second tortfeasor liable for the increment in harm (*Landes/Posner,* JLS 1980¸ 517, 548).

[132] *Landes/Posner,* JLS 1980¸ 517, 524 f.

administrative costs due to additional litigation between tortfeasors.[133] Yet they recognize that the argument no longer holds if one changes some assumptions, for instance, if care is stochastic (ie investments do not prevent the harm with certainty), if the liability standard is unified for all tortfeasors but there are heterogeneous care costs, or if judges err in applying the standard.[134] Kornhauser and Reversz extend the analysis and show, for example, that a no-contribution rule may be equally efficient as a contribution rule under negligence, but is inefficient under strict liability.[135] Mark Grady criticises these perspectives, arguing they are a byproduct of orthodox economics that neglects 'negligence dumping,' when one individual is held negligent for failing to correct the negligence of another.[136]

76 Finally, in a setting closer to AI defects, Jakob and Lovat consider a case where a technology provider passes a defective component on to an industrial operator. They show that the optimal liability allocation rule depends on whether the provider has market power and, in case of a competitive market, parameters like the ratio of efficiency to the cost of prevention and individual relative wealth.[137]

77 Hence, L&E does not offer one recommendation for multiple tortfeasors but rather a broad range of insights that are conditional on various factors.[138]

B. Ambiguous causality

78 While choosing a division rule for multiple tortfeasors is already a daunting task, the situation is even more complex when there is ambiguity regarding causality. Specifically, when it is known that there are multiple tortfeasors but it is unclear which one caused the accident, the L&E debate deepens.

79 Some have argued that the combination of multiple tortfeasors and ambiguity as to who caused the harm prevents an efficient outcome. For instance, Shavell argues that when tortfeasors (i) act sequentially and independently and (ii) are subject to a strict liability standard, then there is simply no division of liability that yields efficiency.[139] However, others have argued that many of the issues can be solved by using a proportional liability rule, where each tortfeasor's

[133] *Landes/Posner*, JLS 1980¸ 517, 529.
[134] *Landes/Posner,* JLS 1980¸ 517, 525.
[135] *Kornhauser/Revesz*, YLJ 1989, 831, 834.
[136] See generally *MF Grady,* Multiple tortfeasors and the economy of prevention, JLS 1990, 653.
[137] *Jacob/Lovat*, CKLR 2016¸659, 679.
[138] Cf. *Jacob/Lovat,* CKLR 2016¸659, 661 f.
[139] *S Shavell,* Economic Analysis of Accident Law (1st edn. 2007) 164-65; *M Gilboa/Y Kaplan/R Sarel*, Climate Change as Unjust Enrichment, Georgetown Law Journal (GLJ) 2023¸ 1039, 1091 fn. 328.

portion is equal to the probability in causation attached to their act.[140] Thus, overall there is no clear consensus on how to deal with ambiguous causality when multiple tortfeasors are potentially involved.

III. Application to AI defects under the PLD

80 From a L&E perspective, definitions are, per se, largely unimportant. It does not matter whether AI is a product or not, how a defect is defined, and how consumer expectations are precisely measured. Instead, what matters is the bottom line: given that we adopt certain filters through definitions and liability standards, who ends up being liable and when.

A. AI products, market failures, and product liability as a tool

81 To kick off the discussion, one may ask whether the PLD only applies to AI defects that actually generate negative externalities. Consumers' expectations are not a natural proxy for negative effects on third parties, so there may be cases where the PLD is over-inclusive or under-inclusive. Furthermore, one may wonder whether tort law liability is the most efficient legal tool to deal with AI defects. As I have analysed this in details in previous work, let me mention briefly that if liability is designed properly, it *can* serve as the most effective tool for certain types of harms and victims (eg those who have sufficient incentive to sue) but not others.[141] For example, a person who was hit by a defective autonomous vehicle and lost a limb clearly has an incentive to sue but a person who dislikes the fact that history books may accidentally cite fake news generated by AI has little incentive to sue.

82 As the PLD exists alongside EU regulations, including the AI Act, the next question is whether regulation plus liability is the optimal solution. For example, some scholars have discussed the possibility of adopting 'data taxes' or 'robot taxes,'[142] to deal with AI, due to its potential negative externalities. If the problem is simply too much use of AI, such 'Pigouvian taxes' may be a simple way to discourage over-use. However, taxes often yield other economic distortions and AI defects occur not only due to overuse but also due to harmful use. As another example, in cases where liability does not work well for AI defects (eg for highly dispersed harms with a low incentive to sue), perhaps unjust enrichment law can be more effective.[143] Nevertheless, given that the

---

[140] *R Young/M Faure/P Fenn*, Multiple tortfeasors: An economic analysis, Review of Law & Economics 2007, 111, 130.

[141] *Sarel*, UCLJ 2023, 115, 166.

[142] *R Kovacev*, A taxing dilemma: robot taxes and the challenges of effective taxation of AI, automation and robotics in the fourth industrial revolution, Ohio State Technology Law Journal 2020¸182; *O Ben-Shahar*, Data pollution, Journal of Legal Analysis 2019, 109.

[143] See *Gilboa/Kaplan/Sarel,* GLJ 2023¸ 1039, for a parallel discussion in the case of climate change; *A Gordon-Tapiero/Y Kaplan*, Unjust enrichment by algorithm, George Washington Law Review 2024, 305; *Y Hu*, Unjust Enrichment Law and AI, in: *E Lim/P*

PLD seems efficient for at least some cases, let us focus on its details with respect to AI defects.

B. The PLD's liability regime

83 When AI is concerned, we are worried about both the level of care—whether economic operators and victims take precautions—and the activity levels—how much AI is provided and used. Thus, we must first check which liability standard the PLD applies. Technically speaking, it applies a strict liability regime, as an economic operator is liable irrespective of whether they are at-fault. However, since defects may hint at some sort of misconduct (assuming that proper care would not yield a defective product) one can think of product liability as a sort of 'middle ground' between negligence and strict liability.[144] Then again, presumptions of a defect bring the PLD much closer to full-blown strict liability. Whether the presumptions will be triggered often or seldom is an empirical question, but let us consider a few examples.

84 First, consider again the *Garcia* case, and suppose it would have been filed in the EU and not the US. Assuming the plaintiffs meet the minimal evidentiary standard, they could ask the defendants for disclosure under the PLD. But will the defendant comply? A generative AI algorithm has some black box features and disclosing the data it was trained on or the exact weights used would likely lead to the disclosure of trade secrets. The PLD tries to balance that by asking member states to consider the legitimate interests of all parties, including issues of confidential information,[145] but the defendants may well refuse to disclose information, immediately triggering strict liability. But even if the defendants comply, recall that the plaintiff only has to point at a likely defect to trigger the presumption again.

85 Second, consider another case: the death of *Elaine Herzberg*, who got hit by an autonomous Uber vehicle while crossing the road with her bicycle because the vehicle misidentified her.[146] In such a case, the 'obvious malfunction' condition would likely trigger the presumption instead.[147] Some have wondered whether the obvious malfunction presumption is even necessary: if the defect is obvious, can the plaintiff not simply prove it easily? The meaning of "obvious malfunction" is not fully clear, but perhaps one must distinguish between "malfunction" and "defect". For instance, an autonomous vehicle may cause harm with a very low probability ex-ante (and hence meet consumer's expectations) but still malfunction sometimes. The presumption would then

Morgan (eds), The Cambridge Handbook of Private Law and Artificial Intelligence (2023).

[144] See *Buiten,* EJLE 2024, 239, 240.

[145] PLD, art 9(3).

[146] *J Shaw,* Artificial intelligence and ethics, Harvard Magazine 2019¸1.

[147] *Kahn* (fn 76) 8.

mean it is up to the defendant to prove that the harm resulted from a rare case and not an ordinary one.

86 Third, as a hypothetical case, consider an AI algorithm that complies with all safety regulations, involves no trade secrets (ie disclosure is easy), and causes highly unusual harm. In such a case, the defect would need to be proven through the CET and the relevant circumstances. If that succeeds, strict liability again applies. The question is whether this is efficient.

87 There are many aspects in which the PLD 'gets it right.' Suppose one is dealing with a case *a la* the *Herzberg* case, where a single manufacturer who produces an autonomous vehicle entirely in-house causes the death of a pedestrian. If the pedestrian cannot do anything to prevent the accident (unilateral care), the PLD provides the advantages of strict liability, which induces the same level of care and more efficient activity levels than negligence. And if the pedestrian *can* do something (bilateral care) the PLD allows member states to reduce the liability. Similarly, the PLD's joint-and-several liability regime is consistent with L&E's discussions of optimal rules for multiple tortfeasors. Its choice to apply a 'contribution' rule (indemnification across tortfeasors) with a strict liability regime (rather than negligence) also avoids some of the disagreements between scholars on the division of liability.

88 At the same time, the PLD 'gets it wrong' on other fronts. First, the ambiguous CET can cause some incentive problems. For example, firms may be over-deterred because they are unsure precisely what the courts will classify as consistent with consumers' expectations, especially in AI products that continuously adapt and present new challenges. Alternatively, the legal uncertainty might discourage lawsuits, for instance, when risk-averse consumers fear losing the case. If firms anticipate this, it may lead to under-deterrence instead.

89 Second, the combination of strict liability and contributory negligence does not solve the issue of the victim's activity levels.[148] Unless the victim is held liable for the entire social cost it creates by using the AI, an efficient activity level is not guaranteed. The problem gets even worse if AI products also have a negative impact on third parties other than the victim, for instance, by leaving a large carbon footprint that consumers do not internalize.[149] However, it should be noted that some features of the PLD mitigate the problem of excessive

[148] See *Cooter/Ulen* (fn. 118) 204 (table 6.2).

[149] For environmental implications of training AI models, see *A Chien/L Lin/H Nguyen/V Rao/T Sharma/R Wijayawardana,* Reducing the Carbon Impact of Generative AI Inference (today and in 2035) in: Proceedings of the 2nd workshop on sustainable computer systems (2023).

activity by victim, including the restriction of liability to reasonably foreseeable use and the limit on the scope of harms covered.[150]

90 Third, the exemptions given by the PLD do not necessarily reflect the relevant distinctions. For example, the DRD may be inconsistent with the LCIG principle because it gives defendants incentives not to seek out new knowledge,[151] such that even a defendant who can easily expand the objective state of knowledge will refrain from doing so. Otherwise, the DRD also has some advantages, as it may incentivize consumers to either seek out information themselves or to buy relevant insurance.[152]

C. Current state of affairs: procedural and institutional choices

1. Strategic substitutes and complements

91 Are there other EU-level legal mechanisms that serve as strategic substitutes for the PLD as far as AI defects are concerned? The AI Act certainly overlaps with the PLD in its aim to address AI risks, but it uses ex-ante regulation rather than ex-post liability. In that sense, it is more of a complement. In fact, there is some intertwining involved between the PLD and the AI Act: the formulation of the CET that refers to the 'legal requirement' implicitly introduces the AI Act as one channel through which a defect can be identified. Thus, proving non-compliance with the AI Act can, by itself, constitute a defect. Furthermore, recall that non-compliance with union law (including the AI Act) can be used to trigger the presumption of a defect. The complementarity between the AI Act and the PLD can make the directive more efficient. However, as mentioned above, using a regulatory standard to automatically trigger liability is problematic because it may distort deterrence if firms have heterogeneous compliance costs.[153]

92 As a counter-example, the EU's general product safety regulation ('GPSR'),[154] which aims to tackle product safety more broadly, explicitly states that its application should not affect the decision as to liability under national law and

[150] Recall that harm to the product itself, potentially caused by overuse, is not covered. Similarly, harm to property used for professional purposes, where business activity levels seem more relevant, is also not covered (PLD, art 6).

[151] *Sarel*, UCLJ 2023, 115, 171.

[152] *Li/Faure,* JETL 2024, 140, 164.

[153] Specifically, firms with high compliance costs for which the Hand formula would suggest there should be no liability, may suddenly (and inefficiently) comply due to the threat of regulation. Respectively, firms with low compliance costs may stick to the bear minimum if doing so exempts them from both liability and regulation. The exact incentives depend on whether dual (non-)compliance with tort and regulation yield additional cost (savings). See *Sarel*, UCLJ 2023, 115, 133.

[154] Regulation (EU) 2023/988 on general product safety of 10 May 2023.

should not affect the PLD.[155] Thus, the GPSR declares that it complements the PLD without being intertwined with it. Yet the PLD mentions "relevant product safety requirements" as one circumstance that requires consideration when deciding whether there is a defect.[156] Thus, it is unclear whether the GPSR and PLD are truly independent from one another. Otherwise, a full picture of EU legislation may require diving into a variety of additional directives and regulations, including the General Data Protection Rules, Data Act, the 'twin directives,' Digital Services Act, and Digital Markets Act,[157] but as there is currently no AILD in force (after the previous version was withdrawn) the question is which types of cases fall through the cracks.

93 The lack of an AILD may create some inefficiencies. First, without harmonization, there may forum shopping on the part of both plaintiffs (ex-post, strategically filing lawsuits in certain member states) and defendants (ex-ante, providing different services in member states where litigation is more likely). Second, plaintiffs would invest resources in framing their lawsuits around product defects, instead of proving negligence. Consequently, the focus will be about what consumers reasonably expect rather than the fault of a single defendant. Depending on the case at hand, this may be either more or less costly. For instance: in the *Garcia* case, showing that consumers reasonably expect chatbots not to encourage suicide seems easier than proving that the app's designers acted unreasonably ex-ante. Conversely, in the *Herzberg* case, proving the exact degree of safety that the public expects from an autonomous vehicle may be complex, whereas showing that the specific vehicle was insufficiently careful in the incident itself compared to alternatives seems straightforward.

94 Some have pondered whether the AILD is redundant, as competition between member states might be sufficient to ensure the adoption of the most efficient rules.[158] Yet competition may also yield the opposite outcome here: a "race to the bottom" in order to attract AI companies, yielding low care incentives.[159]

---

[155] Reg (EU) 2023/988, art 43 (referring to the old PLD, which presumably should be interpreted as applying also to the new PLD).

[156] PLD, art 7(2)f. See also PLD, rec (34).

[157] Directive (EU) 2019/771 of 20 May 2019 on certain aspects concerning contracts for the sale of goods; Directive (EU) 2019/770 of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services; Regulation (EU) 2023/2854 of 13 December 2023 on harmonised rules on fair access to and use of data (Data Act); Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market For Digital Services (Digital Services Act); Regulation (EU) 2022/1925 of 14 September 2022 on contestable and fair markets in the digital sector (Digital Markets Act); Directive (EU) 2022/2555 of 14 December 2022 on measures for a high common level of cybersecurity across the Union (NIS 2 Directive).

[158] *S Li/M Faure*, RE 2025, 115, 123.

[159] *S Li/M Faure,* RE 2025, 115, 125.

There are other justifications for the AILD as well (eg cross-border externalities and saving on transaction costs in the internal market).[160]

2. The attributes of the courts

95 Analysing judicial institutions' characteristics also reveals several efficiency concerns. If courts lack the expertise necessary for a sophisticated evaluation of AI systems, this yields so-called 'epistemic asymmetry' between courts and technologically sophisticated defendants.[161] While the PLD's presumption mechanisms attempt to address this deficit by shifting burdens in cases of technical complexity, this approach may generate economically suboptimal outcomes. Specifically, the presumption may encourage plaintiffs to strategically characterize cases as technically complex in order to benefit from burden-shifting.

96 Epistemic asymmetry necessitates reliance on expert testimony, raising further institutional competence questions. Courts evaluating opposing expert claims regarding AI systems must differentiate between genuine technical disagreements and strategically framed presentations without possessing the requisite technical foundation for independent evaluation. For instance, in litigation concerning cases like the *Herzberg* case, judicial bodies must assess conflicting expert opinions regarding complex probabilistic decision-making processes and autonomous AI without the technical capacity to independently verify claims. This institutional limitation may result in overreliance on secondary heuristics rather than substantive technical assessment, potentially producing economically inefficient outcomes.

97 The CET presents additional institutional challenges. Courts must objectively determine appropriate safety expectations for novel AI technologies—effectively serving as proxies for consumer sentiment about rapidly evolving technical capabilities. Consequently, judges in different cases may reach very different results. The decentralized structure of EU judicial systems introduces further complexity. Absent a harmonized AILD, member state courts may develop divergent interpretations regarding the PLD's application to AI systems, where some interpret the scope widely and others narrowly. Such institutional fragmentation would generate legal uncertainty for economic operators, potentially resulting in inefficient over-deterrence as operators adjust behaviour to comply with the most stringent potential interpretations. This institutional challenge potentially undermines the PLD's harmonization objectives through potentially inconsistent application of the CET across jurisdictions. Whether this can be mitigated through unified EU court rulings remains to be seen.

---

[160] *S Li/M Faure,* RE 2025, 115, 130.

[161] *J Taipale*, Judges' socio-technical review of contested expertise, Social Studies of Science 2019, 310, 310.

98 Temporal limitations further constrain judicial capacity. Given the rapid evolution of AI technologies, substantial temporal gaps between alleged defect occurrence, litigation initiation, and final resolution may render judicial assessments obsolete before judgment. This temporal disconnect is particularly pronounced for continuously learning systems that substantially modify their operation post-deployment.

3. The burden of proof

99 The PLD's burden allocation mechanisms warrant careful examination through a L&E lens. Recall that there are different paths to presume a defect, including (i) a disclosure-based presumption (when the defendant fails to disclose relevant information), (ii) a technical-complexity presumption (when the defendant faces excessive difficulties to prove the defect), (iii) a regulatory-compliance presumption (when the product does not comply with safety regulation), and (iv) an obvious malfunction presumption (when the harm is typical of a defect). I consider each briefly in turn.

100 The disclosure-based presumption functions primarily as an information-forcing device within asymmetric information contexts. Economic operators typically possess superior access to information regarding product design, development processes, and operational characteristics—particularly for proprietary AI systems. This presumption theoretically enhances allocative efficiency by incentivising voluntary information disclosure. In other words, it is an application of the LCIG principle. However, AI products present unique complexities regarding disclosure incentives, as production of the requisite evidence may necessitate revealing algorithmically embedded intellectual property or commercially sensitive training methodologies (as mentioned above). The PLD's attempt to balance these competing interests through judicial discretion regarding 'legitimate interests' introduces substantial uncertainty into the liability calculus, potentially generating strategic non-disclosure incentives in litigation contexts where disclosure costs exceed expected liability costs.

101 The technical-complexity based presumption addresses a potential under-deterrence problem that would arise if valid claims were systematically defeated by practical evidentiary barriers. However, this presumption creates a complex incentive structure for AI developers. On the one hand, it may encourage developers to be keep things simple and transparent, thereby circumventing claims of technical complexity. On the other hands, too much simplicity may undermine the quality of the AI product, which can require certain degree of complexity to satisfy the consumers' needs. Thus, there is a concern of a 'chilling effect', where AI developers are deterred from producing efficient systems simply because doing so raises the chance of litigation.

102 The regulatory non-compliance presumption establishes direct linkage between safety regulations (eg the AI Act) and liability determinations. While this

enhances regulatory compliance incentives, it may simultaneously generate inefficient precautionary measures.[162] As explained above, heterogeneous firms facing uniform compliance requirements but deriving variable benefits from particular AI functionalities may respond inefficiently. Specifically, automatic liability presumptions triggered by regulatory non-compliance may induce excessive precautionary investments by some economic operators while allowing others to undertake insufficient precautionary measures relative to the social optimum.

103 The 'obvious malfunction' presumption extends traditional *res ipsa loquitur* principles to AI products, but encounters conceptual difficulties when applied to probabilistic decision-making systems. AI systems—particularly those operating in stochastic environments with continuous learning capabilities—exhibit substantially different failure characteristics. For example, an autonomous vehicle executing an unexpected manoeuvre might be implementing an optimal response to unusual environmental conditions rather than manifesting a defect. The PLD's application of this traditional presumption to probabilistic AI decision-making risks systematic misclassification of non-defective products as defective, potentially resulting in different incentive problems, such as (i) discouraging efficient care by reducing the differential between the payoffs from good and bad behaviour, (ii) inducing excessive care to avoid extremely rare (but obvious) malfunctions; or (iii) chill innovation because the expected liability is simply too high given the uncertainty.. Beyond the presumptions of a defect, recall that the PLD also adopts a presumption for the causal link in cases where the harm is (stereo)typical of a defect. As I already pointed out some law and economics implications of such presumptions in my previous work,[163] let me provide one additional issue from behavioural law & economics: focusing on typical harm may trigger the so-called 'representativeness heuristic,' where people overestimate the likelihood of typical events that occur only conditional on another event, while ignoring the base rate of that other event.[164] For example, suppose that AI is only rarely defective, but in the very few cases that a defect does exist, there is some very typical harm (eg loss of data). The PLD tells us that if we observe such harm, we should presume a causal link *if* it has already been established that there is, in fact, a defect. That seems fine, if applied at face value. But the representativeness heuristic increases the probability of mistakenly concluding that a defect exists in the first place. Thus, the presumption may exacerbate the consequences of the heuristic, biasing the result. However, whether that will happen often is an empirical question and beyond the scope here.

104 The burden of proof framework further interacts with the various defences, most notably the DRD. For continuously learning AI systems, identifying the exact boundaries of scientific knowledge at particular temporal reference points

[162] For additional discussions, see *Li/Faure,* JETL 2024, 140, 168 f.
[163] *Sarel*, UCLJ 2023, 115, 171.
[164] See eg *C Guthrie/JJ Rachlinski/AJ Wistrich*, Inside the judicial mind, CLR 2000, 805.

is problematic. While defendants bear the evidentiary burden regarding this defence, judicial bodies often lack sufficient technical expertise to properly evaluate such claims. The DRD creates a tension in information-gathering incentives even under strict liability: it enables producers to strategically avoid acquiring information about risks so they can later claim such risks were objectively undiscoverable,[165] but it also prevents over-deterrence by exempting producers from liability for genuinely unforeseeable risks that no amount of reasonable information-gathering could have revealed. This creates a risk that the DRD becomes either excessively accessible (generating under-deterrence by allowing producers to escape liability through strategic ignorance) or practically impossible to establish (creating over-deterrence when producers face liability even for truly undiscoverable risks) depending on judicial interpretation.

Part VI. Conclusion

105 The analysis reveals that while the EU's new Product Liability Directive represents a step forward in addressing AI-related harms, it introduces significant interpretative challenges that may undermine its effectiveness. The PLD's expansion to AI products and its implementation of strict liability principles partly aligns with economic efficiency, particularly through joint-and-several liability provisions and contribution rules that distribute responsibility among multiple economic operators. However, the PLD's reliance on the consumer expectation test proves problematic for autonomous and self-learning AI systems, where reasonable safety expectations are inherently difficult to determine. The complicated interplay between burden-shifting mechanisms, technical complexity presumptions, and various defences makes it difficult to determine whether liability leads to systematic over-deterrence or under-deterrence. The withdrawal of the complementary AI Liability Directive further complicates the regulatory landscape, potentially creating inefficiencies through inconsistent application across member states and strategic forum shopping.

106 This fragmented approach, combined with courts' institutional limitations in evaluating AI systems, suggests some promising avenues for future research. Empirical studies examining how courts apply the PLD's provisions to AI products across different EU jurisdictions would provide valuable insights and experimental research can help test alternative liability frameworks, potentially informing broader discussions about liability regimes for emerging technologies.

[165] Cf *Sarel*, UCLJ 2023, 115, 146; ibid 171.