Technology and Regulation

The Proposed EU Artificial Intelligence Liability Directive

Does/Will Its Content Reflect Its Ambition?

Shu Li*, Béatrice Schütte**

Artificial Intelligence Liability Directive, harmonisation, damage liability, AI Act, disclosure of evidence, causal link, European tort law

On 28 September 2022, the European Commission released its long-awaited proposal for an Artificial Intelligence Liability Directive (AILD). In contrast to the high expectations on providing a harmonised liability framework for the damage caused by AI systems, the proposed AILD only proposes minimum harmonised procedural rules to facilitate evidence disclosure and alleviate the burden of proof undertaken by claimants. This article provides a comprehensive analysis of the proposed AILD and points out the problems when implementing the proposed rules. This article argues that the AILD may never reach its full potential as its name indicates. The fragmentation among Member States regarding the substantive matters may preclude the AILD from moving a step further for harmonising substantial issues. While a comprehensive risk regulation (the EU AI Act) must be followed by an effective remedy mechanism, the proposed AILD will not fill this gap in the short run.

li@law.eur.nl
beatrice.schutte@helsinki.fi

1. Introduction

On 13 March 2024, the long-awaited EU AI Act was finally approved by the European Parliament, marking the birth of the first horizontal regulation on artificial intelligence (AI) in the world.[1] The EU AI Act applies a risk-based approach to regulate different types of AI systems. AI systems that can pose dangers to safety and fundamental rights will in principle be allowed to circulate in the market unless the risks are unacceptable (e.g., if AI is used for social scoring). Although the AI Act lays down the essential requirements that every high-risk AI system must meet before it can be put on the market, it cannot guarantee that harmful consequences will not occur. It is therefore equally important to adapt *ex post* rules to effectively address the harm caused by AI.

Traditionally, tort liability has been considered a crucial means of dealing with harmful consequences. Not only can a liability rule provide tortfeasors with an incentive to optimise their level of care and level of activity, but it can also serve as an important tool to compensate injured persons. Whenever harm occurs, an injured person may have two options under the liability framework.[2] First, they may claim against the producer of the product on the basis of strict liability, subject to the condition that the damage must have been caused by a defect of that product. Second, in most cases, the injured person may refer to fault-based liability. As far as the first avenue for claiming redress is concerned, it must be noted that a harmonised Product Liability Directive has been in place since 1985.[3] In contrast, in all other cases, the damages-liability regime is still fragmented at the domestic level.

The EU authorities have been working on a liability regime for AI in parallel with the AI Regulation. On 28 September 2022, the Commission published two proposals for Directives in order to substantiate its plan of adapting liability rules to the demands of the Digital Age. One proposal[4] aims at modernising the existing European product liability framework, which has been approved by the Parliament quickly with some minor changes on 12 March 2024 ('the revised PLD').[5]

1 European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), available at: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html.

* Shu Li: Erasmus University Rotterdam
** Béatrice Schütte: University of Helsinki; University of Lapland.

Received 11 Jul 2023, Accepted 8 Apr 2024, Published 26 Apr 2024

2 Shu Li, Michael Faure and Katri Havu, 'Liability Rules for AI-Related Harm: Law and Economics Lessons for a European Approach' (2022) 13(4) European Journal of Risk Regulation 618.

3 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 [1985] OJ L 210.

4 Commission, 'Proposal for a Directive of the European Parliament and of the Council on liability for defective products', COM (2022) 495 final.

5 European Parliament legislative resolution of 12 March 2024 on the proposal for a directive of the European Parliament and of the Council on liability for defective products (COM(2022)0495 – C9-0322/2022 – 2022/0302(COD)), available at: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0132_EN.html

The other, the Artificial Intelligence Liability Directive[6] (AILD), adapts non-contractual liability rules to the challenges posed by AI. Since there is no harmonised horizontal regime for damages liability at the EU level yet, the first and most important question to be addressed as regards the AILD is the extent to which harmonised rules should be provided. Will the AILD introduce a set of fully harmonised liability rules for damage caused by AI?

Prior to the era of AI, despite some soft laws made by academics, such as the Principles of European Tort Law[7] and the Draft Common Frame of Reference[8], the Member States never agreed on the harmonisation of the law of damages liability.[9] Consequently, it is reasonable to anticipate that it could be more challenging to establish a harmonised liability regime specifically for the damage caused by AI. So far, only a few harmonised liability provisions can be found in sector-specific EU legislation, such as the General Data Protection Regulation (GDPR)[10] or the Product Liability Directive (PLD). The fact that the Member States differ considerably on the meanings of key concepts (e.g., 'damage', 'fault' and 'causation') and procedural issues (e.g. the burden of proof) explains why the harmonisation of damages liability can be so difficult and why, as some scholars have argued, it may be unnecessary.[11]

Against this background, the Commission opted for a staged approach to the harmonisation of AI liability. As a starting point, the instrument strives to remove the procedural obstacles that injured parties face when claiming compensation for damage caused by AI systems. In the future, based on a targeted review and re-assessment, a harmonised liability rule might be developed to govern the specific case of damage caused by AI systems.[12] This staged approach is considered as an appropriate option due to its proportionality, political feasibility and effectiveness.[13]

Consequently, while the AILD could be a first step towards a harmonised regime on AI liability, there is no intention to enact such rules at this point – the proposal does not contain any basis for the claims of injured parties, nor does it harmonise specific concepts such as 'fault', 'damage' or 'causation'. Instead, as stated in Article 1, it lays down common rules on the disclosure of evidence and the burden of proof and thus focuses on procedural matters. This facilitates access to justice, as required by EU law.[14] It is noteworthy that the current version of the AILD defines it as a *minimum harmonisation* instrument. Member States could thereby maintain or develop rules that are stricter than the ones laid down by the AILD.[15]

This article makes two contributions to the research of AI liability in the EU. Firstly, it critically analyses the rules that are enshrined in the AILD Proposal, pointing out the underlying problems and ambiguities. Secondly, it inquires into the future of the AILD, highlighting challenges and investigating the prospects for the formulation of harmonised rules on AI liability. The structure of the paper is as follows: Section 2 explains the scope of the AILD, pointing out the cases to which the AILD will potentially apply and the grey areas in which ambiguity and controversy may arise. Sections 3 and 4 touch upon the fundamental rules proposed by the AILD from the perspective of evidence disclosure and the presumption of a causal link. Section 5 explains the unclarity and uncertainty of the current proposal and further discusses the potential obstacles and challenges to further harmonisation through the AILD.

2. The applicability of the AILD

Pursuant to Article 1(2) of the AILD, it 'applies to non-contractual fault-based civil law claims for damages, in cases where the damage caused by an AI system'. The scope of the AILD is thereby restricted by three conditions: firstly, the claim must be fault based; secondly, the claim must be non-contractual, that is, it must not arise from a contractual relationship; thirdly, the damage in dispute must have been caused by an AI system. These conditions strictly restrict the applicability of the AILD, making sure that it will not be too far-reaching. This section will critically discuss these three conditions and underline the potential challenges for implementation in practice.

2.1. 'Fault-based claim' as the first condition

The first condition for applying the AILD is that it only applies to *fault-based civil liability claims*. Claims that are based on strict liability are thus excluded. This is based on the assumption that claimants face fewer procedural obstacles in claims that are based on strict liability as they do not need to prove fault. For instance, in Germany, the keeper of a vehicle is subject to strict liability in the case of an accident.[16] Assuming that an autonomous vehicle is driven by AI systems, the keeper would not be able to combine the alleviated procedural measures set out in the AILD, such as those on the disclosure of evidence, with the strict liability rule that is set out in § 7 of the German Road Traffic Act.

However, it should be noted that, most likely, the Road Traffic Act will not preclude the (separate) application of the AILD. Accordingly, an injured party may well base its claim against the keeper on the basic fault-based tort liability rule under § 823 of the German Civil Code in combination with the AILD. In other cases, when the claim is against the manufacturer of a product or the processor of personal data and when it is based on strict liability, the rules that are laid down by the AILD will also be inapplicable unless the injured person opts for fault-based rules.[17]

Therefore, the actual applicability of the AILD will largely depend on the claimant's strategy: the AILD can, in principle, be applicable in all cases, as long as the claimant prefers to base their claims on fault-

6 Commission, '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.
7 European Group on Tort Law, Principles of European Tort Law: Text and Commentary (Springer 2005).
8 Christian von Bar et al. (eds.), Principles, Definitions and Model Rules of European Private Law, Draft Common Frame of Reference (Sellier 2009).
9 Commission, 'Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics' COM(2020) 64 final', 12.
10 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119.
11 Michael Faure, 'The harmonisation of EU tort law: a law and economics analysis' in Paula Giliker (ed), Research Handbook on EU Tort Law (Edward Elgar 2017).
12 Commission, 'Impact Assessment Report Accompanying the document Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence' SWD(2022) 319 final, 39.
13 Commission (n 12), 60-61.
14 Charter of Fundamental Rights of the European Union, art 47.
15 AILD (n 6), rec 14.
16 See § 7 StVG (German Road Traffic Act).
17 Shu Li, 'Compensation for non-material damage under Article 82 GDPR: A review of Case C-300/21' (2023) 30(3) Maastricht Journal of European and Comparative Law 335.

based liability. This situation can happen given that the requirements for strict liability are significantly more difficult to meet (e.g., proof of defect under product liability law) than those of a fault-liability provision. In addition, in the situations where the scope of the recoverable damage is more limited under a rule of strict liability than under a fault-based liability provision, the claimant may choose the latter option to claim their losses before the court. For example, in the case where a person suffered non-material damage resulting from a defective product, they may have to recover non-material harm by resorting to the applicable national damages-liability laws that are fault based, as strict product liability may not grant compensation such harm.[18] This is also a typical scenario that claimants can benefit from the evidence disclosure and alleviations of the burden of proof that are enshrined in the AILD. Finally, it is also important to remember that, unlike the PLD, the AILD is not a consumer protection law. Consequently, the AILD is also applicable to injured parties who are acting in a professional capacity and those legal entities, as long as they are suitable claimants according to national law.

2.2. 'Non-contractual claim' as the second condition

Furthermore, the AILD does not apply to contractual liability. In practice, contractual liability and tort liability may co-exist when a harmful consequence occurs. One question that arises often is whether the existence of a contractual relationship between the claimant and the defendant precludes the applicability of tort liability. Member States provide different answers to this question in their private laws. As we will argue in the following section, this divergence can also influence the applicability of the AILD in various Member States.

Some Member States, such as France, have adopted the so-called *non-cumul* principle.[19] It means that the applicability of contract law will preclude the injured party from invoking tort law provisions. Once there is a contractual relationship between, for example, the provider of an AI system and an injured party, the latter can *only* claim damages on the basis of contractual liability. Consequently, in the countries that apply the *non-cumul* principle, it will only be possible to utilise the AILD when there is no contractual relationship between the claimant and the defendant. Therefore, the applicability of the AILD in such countries will be restricted. Without any further clarification, the relevant AI system provider or user will, for instance, not be obliged to provide the evidence at their disposal to the claimant.

In contrast, in Member States that have not implemented the *non-cumul* principle, such as Germany, a claimant can choose to base their claim on either contract law or tort law, depending on which approach can best remedy their losses.[20] For example, a claimant who has suffered pure economic loss can rely on contractual liability for remedy, since tort law is reluctant to protect economic benefits. In comparison, claimants may rely on tort liability for the restitution of other types of harm and they can thereby benefit from the AILD. Viewed from this angle, the AILD will play an important strategic role in countries that have not implemented the *non-cumul* principle, where it will provide claimants with enhanced flexibility.

The AILD, in its current version, does not address the relationship between contract law and tort law, which is not surprising because it is an instrument of minimum harmonisation. Injured parties might be at a significant disadvantage in the Member States that have adopted the *non-cumul* principle. In any event, the final interpretation of the question about the border between contract and tort for the purpose of the application of the AILD will have to be provided by the Court of Justice of the European Union (CJEU). It will also be up to the CJEU to decide whether upholding the *non-cumul* principle is in accordance with EU law in this particular context.

2.3. 'Damage caused by an AI System' as the third condition

The AILD will not apply to all damage that is *related* to an AI system; only damage *caused* by an AI system will trigger its applicability. Recital (15) contains an interpretation of the notion of 'damage caused by an AI system'. It reads as follows: 'this Directive should only cover claims for damages when the damage is caused by an output *or* the failure to produce an output by an AI system *through* the fault of a person'. In this regard, the AILD will apply only if a party's fault has an impact on the output of an AI system and this output generates damage.

In practice, the fault of two parties may potentially influence the output of an AI system. Firstly, fault during the manufacturing and design phases can directly give rise to harmful output from an AI system. In this case, although the injured person has the option of claiming the damage on the basis of strict product liability, as noted previously, they may also choose fault-based liability for certain reasons, such as their interests not being covered by the product liability regime. Secondly, deployers may also influence the output of an AI system. Deployers are normally the ones who decide the use of an AI system for certain purposes and supervise its operation. They are obliged to ensure that the AI system is updated in a timely manner and that it is in the proper condition for the work that it has been delegated to it. It is noteworthy that although non-professional deployers will not be covered by the EU AI Act, the damage that they cause will be covered by the liability regime. Therefore, the AILD will apply to the damage caused by non-professionals.

The AILD also explains the extent to which it is *not* applicable. Recital 15 further states, '[t]here is no need to cover liability claims when the damage is caused by a human assessment followed by a human act or omission, while the AI system only provided information or advice which was taken into account by the relevant human actor'. In this regard, the AILD should not be relevant to a claim for damage which is caused by the faulty interpretation of the output of a specific AI system. The fault, in that case, does not affect the output of the system, and it is considered to interrupt the causal chain between output and damage.[21] Therefore, the claim for wrong assessment against certain *deployers* (e.g., doctors who use intelligent diagnostic tools and drivers of autonomous vehicles) are likely excluded from the scope of the AILD. The basic idea behind this exclusion is that since it is a human act or omission that occurs after receipt of the output of the AI that contributes to the harmful consequence, there is no AI black-box

18 Gerhard Wagner, 'Liability Rules for the Digital Age: Aiming for the Brussels Effect' (2023) 13(3) Journal of European Tort Law 191, 236-237.

19 See e.g. Jean-Sébastien Borghetti, 'The borderlines of tort law in France' in Miquel Martin-Casals (ed), The Borderlines of Tort Law (Intersentia, 2019), 136.

20 The only limitation is that alleviations to liability that the parties agreed on in the contract will have an influence on tort liability. See e.g. Andreas Spickhoff, 'Privatrechtsdogmatik und Deliktsrecht: Das Deliktsrecht in der Konkurrenz zum Vertragsrecht im Spiegel von Kollisionsrecht und Gerichtsständen Privatrechtsdogmatik Im 21. Jahrhundert' in Hans Christoph Grigoleit and Jens Petersen (eds), Festschrift Für Claus-Wilhelm Canaris Zum 80. Geburtstag (Walter de Gruyter GmbH, 2017), 552.

21 Philipp Hacker, 'The European AI Liability Directives--Critique of a Half-Hearted Approach and Lessons for the Future' (2023) 51 Computer Law & Security Review 1.

problem and, thus, there is no need for establishing new rules that alleviate the burden of proof.

In practice, considering the fact that many applications require human-AI interaction, an AI system that makes an inaccurate inference or provides inaccurate advice and the human actors who assess that inference or piece of advice could be thought to contribute jointly to the harmful consequence. In that case, it is likely that the injured person will claim against the different parties on different bases. For example, a patient may claim damages on five legal bases: (a) against the provider of an intelligent diagnostic tool on the basis of strict product liability, (b) against the provider on the basis of fault-based liability, (c) against the hospital on the basis of fault (e.g., lack of supervision), (d) against the doctor who made the erroneous assessment by drawing on AI-generated advice and (e) against the doctor who used the AI device wrongly and thereby received inaccurate advice accordingly. It is likely that the provider, the hospital and the doctor are jointly and severally liable for the damage. However, since the AILD will only apply to fault-based liability and the fault must have impact on the *output* of the AI system, only cases (b), (c) and (e) will be covered by the AILD.

The discussion in this section indicates that the applicability of the AILD is very largely dependent on the concept of causation. This reflects that the applicability of the AILD shall be proportionate. The causation can be established only to the extent that a party's fault has influenced the output of an AI system. The issue that remains, however, is that different Member States may present different criteria to decide the causation in this context. Since causation is a substantive matter that the AILD will not harmonise, we can expect that the applicability of the AILD could be quite different among the Member States.

3. The disclosure of evidence for high-risk AI systems

The first measure proposed by the AILD to alleviate the burden of proof is that Member States shall ensure that (potential) claimants can access relevant evidence easily. Article 1(1)(a) stipulates, 'this Directive lays down common rules on the disclosure of evidence on high-risk artificial intelligence (AI) systems to enable a claimant to substantiate a non-contractual fault-based civil law claim for damages'.[22] This section critically analyses the mechanism of evidence disclosure under the AILD and points out the scenarios in which it might not be as effective as the legislator expected.

3.1. Evidence disclosure as a continuity of the requirement from the draft AI Act

As a risk regulation instrument, the AI Act requires the providers and deployers of high-risk AI systems to comply with numerous regulatory requirements before a system can be placed on the market. One of the essential requirements is high-risk AI systems should be designed and developed with a capability of enabling automatic recordings of events, i.e., 'logs', throughout their lifecycle'.[23] Those who deploy high-risk AI systems are required to make sure that the logging capability is enabled during operations.

The record-keeping requirement that the AI Act lays down for providers and deployers only addresses the problem of the adequate storage of information. It does not guarantee that safety-relevant information of this kind will ultimately be visible and accessible to ordinary individuals and entities. The fact that such information can be technically inaccessible and unreadable to victims or protected as a trade secret or by property rights could significantly hinder victims' attempts to collect information that supports their claim. As a result, whenever harm occurs, specific *ex post* legal instruments must be in place to ensure that the victims can access the relevant information and evidence easily when certain conditions are met. The AILD can thus contribute to the effort to plug the *ex post* information gap by providing rules on evidence disclosure.

One question that remains unclear is whether the Member States can extend evidence disclosure to other AI systems that do not pose a high risk. It seems that the AILD does not preclude this possibility. Since the AILD was proposed as minimum harmonisation legislation, it does not prevent the Member States from passing stricter rules. However, extending evidence disclosure to the AI systems other than high-risk ones may not be consistent with the risk-based approach laid down by the EU AI Act. As noted previously, only high-risk AI systems are subject to the record-keeping requirement from the AI Act. If the Member States are allowed to extend evidence disclosure to providers of other AI systems, their providers would become subject to excessive incentives to take *ex ante* measures (e.g., to maintain logs) in order to avoid problems *ex post*. This would simply increase regulatory burdens for other AI systems, which would run counter to the principle of proportionality and risk-based approach that are set out in the AI Act.

3.2. Unwinding the requirements for evidence disclosure under the AILD

The AILD in Article 3(1) requires different procedures for evidence disclosure, depending on the requesting party (the claimant[24] or the potential claimant[25]) and on the target of the request (the defendant or a third party). This subsection provides a critical analysis of the three situations that the AILD has differentiated in light of the disclosure of evidence.

3.2.1. Situation A: The claimant requests evidence from the defendant

The first situation in which a request of evidence disclosure can occur is when a *claimant* asks a national court to order the disclosure of evidence that is at the disposal of providers or deployers who are already *defendant(s)* in the process of litigation. This is the most common case in practice. For example, a person who has suffered personal injury while using an autonomous lawnmower may claim damages from its deployer if the latter has failed to provide updates, thus causing the damage. In this case, the relevant evidence to support the claim is likely at the disposal of the deployer.

When considering the request for disclosure, the court must balance the interests of the different parties that are involved. According to Article 3(4) of the AILD, 'the national courts should limit the disclosure of evidence to the extent necessary and proportionate to the claim for damages'.[26] To that end, the national courts must

22 AILD (n 6), art 1(1)(a).
23 AILD (n 6), art 12.

24 According to Art 2(6), the term 'claimant' refers to a natural or legal person that has already filed a case before the court because they were injured by the output of an AI system, due to the fact that they succeeded or were subrogated to the right of an injured person, or because they are acting on behalf of injured persons.
25 According to Art 2(7), the term 'potential claimant' refers to a natural or legal person, who 'is considering but has not yet brought a claim for damages'.
26 AILD (n 6), art 3(4), para 1.

consider the legitimate interests of all parties, which are related not only to particular public interests, such as national security, but also to private ones, such as the protection of trade secrets and confidential information.[27] In particular, when confidential information is involved, the national courts shall have effective means and be empowered, upon the request of the relevant parties, to take measures that are necessary to preserve its confidentiality.[28] Member States must also ensure that parties that are ordered to disclose evidence are provided with 'appropriate procedural remedies' to react to such orders.[29] However, the proposal does not indicate which remedies are considered 'appropriate'. Such remedies would have to be created accordingly. Again, the CJEU has the final say on whether a remedy is appropriate.

3.2.2. Situation B: The claimant requests evidence from a third party

In many cases, the claimant may find that it is not the defendants that hold the necessary evidence but other providers or deployers that are not parties to the litigation. These parties, not being defendants, are likely not at fault for the damage in dispute. However, the claimants may find that these parties hold crucial evidence for proving the fault of the defendants. For example, a patient may sue a hospital because it failed to supervise AI-driven diagnostic tools appropriately and later discover that key evidence, such as logs, is held by the provider of the diagnostic tools. In this scenario, the patient, as the claimant in an existing lawsuit, may need access to the logs.

The AILD covers this situation. The wording of Article 3 (1) and (2) of the AILD indicates that evidence disclosure can also be requested from third parties. Nevertheless, if compared to Situation A, evidence disclosure from third parties comes with additional requirements. According to Article 3(2), before they can request disclosure from a third party, the claimant must have undertaken all proportionate efforts to obtain that evidence from the defendant. In other words, the third party may decline such a request by arguing that the claimant can obtain that evidence from the defendant with a reasonable effort.

3.2.3. Situation C: The potential claimant requests the disclosure of evidence

Finally, according to the AILD, a *potential claimant* can request the provider or deployer to disclose evidence prior to litigation. This can reduce unnecessary litigation and trial costs.[30] Nevertheless, unless the right is subject to reasonable conditions, there is a risk of abuse. Therefore, compared with the aforementioned two situations, disclosing evidence initiated by a potential claimant is subject to more restrictions.

The AILD sets out two conditions that must be met before a national court can proceed with the disclosure procedure. Firstly, the potential claimant can initiate a request for evidence disclosure *only if* they have already asked the relevant provider or deployer to disclose the evidence at their disposal and had their request rejected.[31] In addition, the potential claimants must present further facts and evidence that should be considered sufficient to support the 'plausibility' of a claim for damages when they request evidence disclosure.[32] What constitutes the plausibility of a claim for damages is again expected to be interpreted by the CJEU when a case is referred to it.

3.2.4. Discussion

The foregoing shows that the AILD sets out different requirements for the disclosure of evidence under various conditions. If the request is initiated by potential claimants or against target third parties, the requirements will be more stringent. The analysis in this part also indicates that national courts have a lot of discretion when evaluating the request for evidence disclosure. In any event, unless the existing ambiguities are addressed in the course of the legislative process, which is ongoing, the CJEU will have to clarify the requirements, particularly in relation to necessity, proportionality and the balancing of the interests of the parties.

3.3. The consequence of failing to disclose evidence: the presumption of non-compliance

If the order of a court to disclose evidence is not followed, Article 3(5) of the AILD stipulates that the national court must *presume* that the defendant *has not complied with* the relevant duty of care under EU or national law, that is, that the claimant has proven what they intended to prove with the evidence that they sought. The defendant can rebut that presumption by submitting evidence to the contrary effect.[33]

3.3.1. Non-compliance ≠ fault

The key issue here is whether the wording of Article 3(5) is intended to link the failure to disclose evidence directly with a finding of fault. It is admitted that the development in national laws of Member States shows a certain tendency towards a more objective concept of fault, so that there is no longer a strong need for a subjective element.[34] Nevertheless, the violation of a duty of care does not automatically equal fault at the national level. Because of differences in private-law traditions, some Member States may still require a subjective element (e.g. culpability) when deciding the fault.[35]

Therefore, the question of whether non-compliance can be regarded as a way of establishing fault should only be answered in line with the rules of national law.[36] On the one hand, if non-compliance is considered sufficient for certain cases in a certain Member State, then the mere failure to comply with an order to disclose evidence would be treated as equivalent to fault. On the other hand, according to the law of a given Member State, if objective non-compliance does not necessarily lead to the establishment of fault, the claimant would still need to prove that the other conditions for such a finding have been fulfilled.

It is noted that even under the revised PLD a non-compliance does not necessarily lead to a presumption of defectiveness. According to the new provisions, only the non-compliance with 'mandatory' safety requirements can be presumed as a defectiveness. Mandatory safety requirements are normally the ones that have been agreed among the Member States, so it makes sense to link it with defectiveness. In contrast to the revised PLD, the AILD does not distinguish a non-compliance of mandatory safety requirements from others.

27 AILD (n 6), art 3(4), para 2.
28 AILD (n 6), rec (20) and art 3 (4), para 3.
29 AILD (n 6), art 3 (4), para 4.
30 AILD (n 6), rec (17).
31 AILD (n 6), art 3(1), para 1.
32 AILD (n 6), art 3(1), para 2.
33 Art 3(5) and Recital (21), AILD (n 6).
34 See e.g. Gerhard Wagner, 'BGB, § 823' in Franz Jürgen Säcker et al., Münchener Kommentar zum Bürgerlichen Gesetzbuch (C.H. Beck, 2015)
35 Jan de Bruyne, Orian Dheu and Charlotte Ducuning, 'The European Commission's approach to extra-contractual liability and AI – An evaluation of the AI liability directive and the revised product liability directive' (2023) 51 Computer Law & Security Review 1.
36 AILD (n 6), rec (14).

Therefore, the consequence of non-compliance should be in line with the national law. To summarise, we can reasonably expect that the effect of the AILD to evidence disclosure will differ considerably across domestic jurisdictions. Nevertheless, as we will show in the next section, despite the difference among domestic tort law regarding the concept of fault, one thing of the presumed non-compliance will be certain: the presumption of non-compliance will be sufficient to establish fault *for the purposes of* the presumption of a causal link.

3.3.2. The effectiveness of the disclosure of evidence

The effectiveness of evidence disclosure remains unclear at the moment. So far, the analysis indicates strongly that the AILD may only facilitate evidence disclosure in situations in which the defendant is also the party that holds the requested evidence (Situation A).[37] Conversely, it is always problematic to oblige natural or legal persons to disclose information if they are third parties to a litigation (Situation B and Situation C). The third party is not subject to any disadvantage if it chooses not to disclose evidence. In this regard, our analysis implies that there is no incentive to boost evidence disclosure in such situations. Considering this deficiency, the effectiveness of the evidence disclosure rules of the AILD can be low in practice.

Another issue that remains ambiguous in the current proposal has to do with the interpretation of the term 'evidence to the contrary', based on which a defendant can rebut the presumption of a causal link successfully. For instance, a defendant may argue that their AI systems are not high-risk ones. They could also argue that, without effective safeguards, the disclosure of evidence may place trade secrets at risk. It remains to be seen whether this provision on rebuttal will work as expected, that is, whether it will facilitate court proceedings in practice, or whether it will become a weapon by which the defendant can escape the presumption of non-compliance. Only the CJEU can clarify the question of what can be used as 'evidence to the contrary' conclusively.

4. The presumption of a causal link in fault-based cases

The opacity of an AI system can make it difficult for a claimant to prove the existence of a chain of causation between the defendant's fault and the output of the system.[38] Several options for lightening the burden of proof in relation to the causal link have been discussed. Those options range between the stringent, such as a total reversal of the burden of proof, and the moderate, such as a (rebuttable) presumption of a causal link.[39] Previously, both the EP Resolution[40] and the EG-NTF Report[41] provided, to some degree and subject to certain conditions, for the reversal of the burden of proof for causation.[42] In the proposed AILD, however, a rebuttable presumption was considered the most proportionate approach.[43] Nevertheless, since the AILD follows a *minimum harmonisation* approach, the Member States can maintain or introduce more stringent rules for the establishment of a causal link.[44]

4.1. The basic conditions for presuming a causal link in fault-based cases

According to Article 4 (1) AILD, national courts must presume the existence of a causal link between the fault of the defendant and the output produced by the AI system if *all* of the three conditions that follow are met: establishment of fault, proof of reasonable likelihood and demonstration of damage.

Firstly, the AILD adopts a lower threshold for establishing fault requirement for the purpose of proving causation. Specifically, if the claimant can prove that the defendant *failed to comply with the duty of care* from Union or national law which is intended to protect them directly from the damage that occurred, said proof would be sufficient to establish fault.[45] According to Article 4(1)(a) AILD, fault through non-compliance can be established in two ways. To start with, the defendant's non-compliance can be demonstrated directly by the claimant. In addition, it can be established in accordance with Article 3(5), whereby the defendant is presumed not to have complied if they have failed to disclose evidence as ordered.

One issue that the AILD does not clarify conclusively has to do with the relationship between the establishment of fault *for the purpose of presuming a causal link* and the establishment of fault *per se*, that is, as an independent element that should be proven by the claimant in all fault-based claims. Since the lower standard for establishing fault for which Article 4(1)(a) provides only applies for the purpose of the presumption of a causal link, the claimant has to prove fault *per se* in accordance with the standards that are set out in national law. Therefore, it seems that the burden of proof sustained by the victim is not alleviated, when it comes to the Member States that a subjective element of proving the fault is required.

Secondly, in order to establish the presumption of a causal link, the claimant must prove that the fault (as per Article 4(1)(a)) was *reasonably likely* to have influenced the output that was produced by the AI system or the failure of the AI system to produce an output.[46] According to Recital (25), a reasonable likelihood of a causal link between fault and damage is not established if the result of the fault did not impact the output of the AI system and therefore did not generate the relevant damage. For example, when it comes to non-compliance with certain requirements from the AI Act, damage may be considered a reasonably likely consequence of non-compliance with the standards on data governance, transparency or human oversight. In contrast, failure to register a certain high-risk AI system, as required by the AI Act, may not be considered because such non-compliance does not have any impact on the functioning and output of the system. In practice, what exactly constitutes a 'reasonable likelihood' would be a matter for the CJEU, and its determination would have to be based on the circumstances of the case before it.[47]

37 AILD (n 6), rec (21).
38 Miriam Buiten, Alexandre de Streel and Martin Peitz, 'The law and economics of AI liability' (2023) 48 Computer Law & Security Review 1.
39 Commission (n 11), 32-33.
40 European Parliament, 'Resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence' (2020) 2020/2014(INL).
41 EG-NTF, 'Report: Liability for Artificial Intelligence and Other Emerging Digital Technologies' (2019) https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf accessed 16 January.
42 EG-NTF (n 41), Key findings [22]-[27]; European Parliament (n 40), art 8(2).
43 Commission (n 12), 74 and 87; AILD (n 6), art 1(1)(b).
44 AILD (n 6), rec (14).
45 AILD (n 6), art 4(1)(a).
46 AILD (n 6), art 4(1)(b).
47 Ljupcho Grozdanovski, 'The Explanation One Needs for the Explanations One Gives: The Necessity of Explainable AI (XAI) for Causal Explanations of AI-related harm-Deconstructing the 'Refuge of Ignorance' in the EU's AI liability Regulation." (2024) JUST-AI Jean Monnet Research Papers https://orbi.uliege.be/handle/2268/312246 accessed 1 February 2024.

Thirdly, the claimant must demonstrate that the output of the AI system or its failure to produce an output gave rise to the damage in dispute.[48] Since the AILD is not intended to provide a harmonised understanding of the concept of 'damage', the scope of that concept will be determined by specific EU or national laws. As a result, this implies that the applicability of the AILD across the Member States could be significantly different.

4.2. Reshaping the conditions with the consideration of the risk profile of AI systems

It should be noted that, unlike the rules on evidence disclosure, which only apply to high-risk AI systems,[49] the rebuttable presumption of a causal link applies to *all* AI systems, regardless of their risk classification, because all such system can be opaque.[50] Nevertheless, the AILD provides for further requirements for and restrictions on the presumption of a causal link, which depend on the risk profile of the AI system.

Firstly, the AILD indicates the fault of the *provider* of a high-risk AI system can be established for the purpose of the causal-link presumption if they fail to comply with requirements and the obligations that are listed in Article 4(2), which include transparency, human oversight and accuracy.[51] Likewise, Article 4(3) stipulates that the fault of the deployer of a high-risk AI system can be established for the purpose of presuming a causal link only if they failed to monitor the AI system in accordance with the accompanying instructions for use or if they failed to ensure that the input data were relevant and representative (insofar as that they exercised control over said data).[52] Therefore, the effectiveness of the AILD will largely rely on the AI Act in light of the essential requirements that it sets out.[53]

Secondly, when the three conditions for presuming a causal link are met, the court may make different decisions, depending on the risk level of the AI system. For claims that concern high-risk AI systems, the court *must* presume a causal link unless 'the defendant demonstrates that sufficient evidence and expertise is reasonably accessible for the claimant to prove the causal link'.[54] For instance, this is the case when the claimant can obtain sufficient evidence and expertise by accessing the documentation and logs of a high-risk AI system.[55] In comparison, for AI systems that are not classified as high risk, the national courts must further determine whether it is 'excessively difficult for the claimant to prove the causal link' before they presume a causal link.[56] One must remember that, according to the AI Act, these systems are not subject to the requirement to keep records. Thus, the documentation in question might not even exist, and the court must decide whether this creates an unsurmountable obstacle for the claimants' efforts to prove the existence of a causal link.

5. Discussion

So far, the analysis implies that it is not easy to address the problems that AI poses for the liability regime. The proposed AILD is a liability framework in the name only; in its nature, it is a procedural law. As far as the substance of liability law is concerned, fragmentation across the Member States will persist. This section raises two specific concerns about the AILD. One is to point out the ambiguity of the text of the current version, and the other is to reflect the difficulty of establishing a full harmonised framework on liability for the damage caused by AI.

5.1. Ambiguity and controversy about the current proposal

Firstly, there is a need to clarify if and to what extent the *non-cumul* principle that applies in some Member States will affect the applicability of the AILD, that is, whether its scope of application in these countries is limited to cases in which there is no contractual relationship between the parties. This question, as well as the border between contract and tort, will have to be addressed by the CJEU. However, many years of legal uncertainty may lie ahead because, in practice, it can take a very long time for a court case in a Member State to give rise to such questions, which would then result in a request for a preliminary ruling from the CJEU.

Secondly, it is not completely clear which actors will be caught by the AILD. As noted previously, the AILD only covers damage 'caused' by an AI system. Recital (15) indicates that causation of this kind can be made out only if a given person's behaviour could have influenced the output of the AI system, which must also be shown to have generated damage. Recital (15) further clarifies that if damage is caused by a faulty human assessment that is based on the advice given by an AI system, that damage is actually the result of human fault. In such a case, there is no need to rely on the AILD rules that alleviate the burden of proof. In practice, however, considering the complexity of human-AI interactions, whether or not a deployer has influenced the output of a system is not always clear cut. For example, in the context of autonomous vehicles, an accident may be 'caused' by the data that is collected from driving activities, which are highly dependent on the habits of drivers (e.g., driving in the evening or in extreme weather conditions). When an accident occurs, it is difficult to determine whether it was caused by the output of the AI system through the fault of a person (i.e., a driver).

Thirdly, limiting the applicability of the AILD to cases in which damage is caused by an AI system can also produce a paradox in practice. A potential claimant who attempts to acquire evidence can only enjoy the evidence-disclosure benefits that are enshrined in the AILD if they can prove that the damage that they suffered was caused by an AI system. However, without access to the relevant evidence, that potential claimant may be unable to prove that their damage was caused by such a system in the first place. Consequently, they would not benefit from the rules of the AILD.

Fourthly, the rules on the disclosure of evidence and the presumption of a causal link are rather vague and complex. As analysed, the AILD equates fault to non-compliance with a duty of care only to the extent

48 AILD (n 6), art 4(1)(c).
49 See the discussion in Section 3.
50 AILD (n 6), rec (28).
51 The relevant requirements in the draft AI Act (n 1) include art 10(2)-(4) (data sets), art 13 (transparency requirements), art 14 (human oversight), arts 15 and 16 (a) (accuracy, robustness and cybersecurity) as well as arts 16(g) and 21 (necessary corrective actions).
52 The term 'non-compliance' refers to the following scenarios in the draft AI Act (n 1): the user fails to monitor the high-risk AI systems in accordance with the instructions (art 29) or the user exposes the high-risk AI system to input data under its control but the data is unnecessary for the intended use (art 29(3)).
53 Marta Ziosi et al., 'The EU AI Liability Directive (AILD): Bridging Information Gaps' (2023) 12(3) European Journal of Law and Technology 1.
54 AILD (n 6), art 4(4).
55 AILD (n 6), rec (27).
56 AILD (n 6), art 4(5) and rec (28).

of presuming a causal link. This equation will not be established in other cases. This is in line with the purpose of the AILD, which is not intended to harmonise substantive matter, such as fault. Nevertheless, further clarification on this issue would be welcome.

Fifthly, it remains to be seen how effective the procedural rules will be in practice. For example, as far as the disclosure of evidence is concerned, the party that possesses the relevant evidence is not always the defendant. This indicates that the (potential) claimant may request a national court to order persons that are not party to the litigation to disclose evidence. Under the current proposal, however, if the third party refuses to disclose the relevant evidence, there are no further means of breaking the deadlock. It is also open to question whether the procedural rules will in fact alleviate the burden of proof that claimants bear. As we showed, while a finding of non-compliance with a duty of care is sufficient to establish the fault that is required for a causal link to be presumed, such a finding does not influence the requirement to prove fault as a separate element of liability. In this regard, since the AILD does not interfere with Member State definitions of the concept of 'fault', the alleviation of the burden of proof that the AILD guarantees can be very limited.

Last but not least, in terms of procedural measures, a substantial number of tasks is assigned to national courts. However, when it comes to the interpretation of EU law, of which the AILD forms part, the CJEU has the last word. For instance, pursuant to Article 3(2) of the AILD, a national court must order evidence to be disclosed only if the claimant has undertaken 'all proportionate attempts' to secure it. The CJEU will have to interpret the phrase 'all proportionate attempts' in order to avoid fragmentation across the Member States.

5.2. Prospects for a harmonised AI liability framework?

The impact assessment and the previous debates (e.g., those reflected in the EP Resolution and EG-NTF Report) point out that, depending on the outcome of a future targeted review, the AILD might turn into a framework that actually harmonises liability for damage caused by AI.[57] In that case, a harmonised set of rules would apply to damage caused by AI systems in all Member States. The discussion in this part will focus on this specific perspective; the argument is that policymakers should address several questions prior to embarking on this second stage.

The first relevant question is whether the AILD should be defined as a Directive that only focuses on liability for safety-relevant harm, like the PLD, or as a Directive that covers all kinds of damage caused by AI systems, such as those with implications for fundamental rights. To date, the AILD has been ambiguous on this point. On the one hand, Article 1(2) clearly stipulates that the AILD applies to all non-contractual fault-based civil law claims, as long as damage is caused by a specific AI system. Therefore, the AILD is not merely a safety-oriented liability regime; instead, it is intended to apply to any harm that results from the use of AI systems. On the other hand, the Explanatory Memorandum indicates that the AILD, together with the AI Act and other Union harmonised safety regulations, is part of a safety package. According to the Memorandum, 'safety and liability are two sides of the same coin: they apply at different moments and reinforce each other'.[58] From this perspective, it seems that the role of the AILD is limited to addressing safety-specific harm. If future policymakers are inclined to consider the AILD an *ex post* liability regime that can cover *any* damage caused by an AI system, the next issue would be that of ensuring that it is aligned with the other sector-specific legislation that also provides for private enforcement in cases of violations of fundamental rights, such as the data protection law and anti-discrimination law at the EU level.

What is more, it is doubtful from the theoretical perspective whether there will be an actual harmonised framework on AI liability. First and foremost, it is difficult to reach an agreement on the set of scenarios to which strict liability applies. Studies have already shown that this could be an impossible task.[59] Member States hold quite different opinions about the criteria for adopting strict liability, and the activities to which it may potentially attach range between the dangerous and the ultrahazardous. This indicates that some countries would be more open than others to adopting strict liability for certain activities. In addition, the harmonisation of recoverable damage could be even more complex than that of liability.[60] The divergence of attitudes toward tort damage, especially for non-material damage, among the Member States cannot be overcome easily because this fragmentation is also linked to cultural and economic differences.[61] So far, there have only been scholarly efforts to identify the common principles of a so-called 'EU tort law', which have not, however, resulted in actual attempts to harmonise that domain.[62]

Furthermore, a harmonised liability regime for AI damage could generate significant tensions with the existing sector-specific legislation that contains liability rules. Take the tension between the AILD and the product liability regime as an example. According to product liability law, producers are subject to strict liability if a defect in their product causes damage. The risk-based approach under the AI liability regime, however, contradicts that approach. Specifically, according to the AILD, AI systems with specific risk profiles are subject to strict liability, but this strict liability does not require the existence of a defect. This indicates that AI systems with a specific risk profile will be subject to more stringent rules than those of the product liability regime. In comparison, the manufacturers of AI systems that are not high risk bear fault-based liability under the AILD. In the latter case, the AILD exempts the providers of certain AI systems from the strict liability that they would bear under the product liability regime. Therefore, the consistency between the AILD and sectoral laws will be a critical issue.

6. Conclusion

The long-awaited proposal for an AILD marks the first step that the EU has taken to address liability for damage caused by AI systems. As the discussion in this article showed, the current version of the AILD aims neither to offer a harmonised liability regime that applies to all Member States nor to disrupt essential concepts whose meaning differs across Member States, such as such as 'fault' and 'damage'. Instead, it only focuses on procedural measures that alleviate some of the obstacles that claimants face when they try to recover for the

57 AILD (n 6), Explanatory Memorandum, 9.
58 AILD (n 6), Explanatory Memorandum, 2.
59 Miquel Martin-Casals, 'The Principles of European Tort Law (PETL) at the beginning of a second decade' in Paula Giliker (ed), Research Handbook on EU Tort Law (Edward Elgar Publishing, 2017), 384.
60 Paula Giliker, 'What do we mean by 'EU' tort law?' in Paula Giliker (ed), Research Handbook on EU Tort Law (Edward Elgar Publishing, 2017), 18.
61 Stephen Sugarman, 'Tort damages for non-economic losses: personal injury' in Mauro Bussani and Anthony Sebok (eds), Comparative Tort Law (Edward Elgar Publishing, 2021), 319-320.
62 Gerhard Wagner, 'The project of harmonizing European tort law' (2005) 42 Common Market Law Review 1269.

damage that they have suffered. In this regard, the AILD only takes a very small step forward. This article revealed how policymakers have striven to deliver the current version of the AILD and critically analysed the rules that are currently proposed for inclusion into the AILD.

This article concludes with concerns for the unclear future of the AILD, which may remain a liability framework in the name only. One obstacle has to do with its historical background. An important reason for the rapid passage of the AI Act was that the pre-existing safety regulations (e.g., the NLF) had already laid down important preparatory measures for harmonisation; consequently, legislators could dedicate more attention to the substantive rules of the Act. In comparison, there was no such common ground prior to the drafting of the AILD. It is unrealistic to harmonise specific liability rules for damage caused by AI systems without reaching an agreement on a common liability framework beforehand. Beyond this historical issue, it is not clear how the harmonised liability regime for AI could be consistent with the other sectoral regimes that already exist. Those regimes already contain provisions that regulate or refer to liability and the recovery of damages. Last but not least, there is a normative question that this article did not aim to address but which can be important prior to harmonisation: does the EU need a harmonised liability regime specifically for damage caused by AI? The law-and-economics literature has shown that the mere fact that Member States have different rules on a given matter does not justify harmonisation.[63] Harmonisation *per se* comes with a cost, and robust justifications, such as transboundary externalities and the prevention of a race to the bottom, should be considered carefully prior to such a development.

63 Michael Faure, 'Product liability and product safety in Europe: harmonization or differentiation' (2000) 54(4) Kyklos 467.

Copyright (c) 2024, Shu Li, Béatrice Schütte

Creative Commons License
This work is licensed under a Creative Commons Attribution-Non-Commercial-NoDerivatives 4.0 International License.