

**AN OVERVIEW OF THE PROPOSAL FOR THE AI LIABILITY
DIRECTIVE IN THE EU**

Marija Ampovska
Goce Delcev University Stip, North Macedonia

E-mail: marija.ampovska@ugd.edu.mk

Introduction

- ▶ It will change our lives by improving healthcare (e.g. making diagnosis more precise, enabling better prevention of diseases), increasing the efficiency of farming, contributing to climate change mitigation and adaptation, improving the efficiency of production systems through predictive maintenance, increasing the security of Europeans, and in many other ways that we can only begin to imagine. At the same time, Artificial Intelligence (AI) entails a number of potential risks, such as opaque decision-making, gender-based or other kinds of discrimination, intrusion in our private lives or being used for criminal purposes.
- ▶ While AI can do much good, including by making products and processes safer, it can also harm. This harm might be both material (safety and health of individuals, including loss of life, property damage) and immaterial (loss of privacy, limitations to the right of freedom of expression, human dignity, discrimination for instance in access to employment), and can relate to a wide variety of risks.
- ▶ The main risks related to the use of AI concern the application of rules designed to protect fundamental rights (including personal data and privacy protection and non-discrimination), as well as safety and liability-related issues.

Background

- ▶ EU citizens and EU businesses' trust in AI technologies is impaired. While European consumers generally consider AI applications potentially useful for their everyday life, such applications are perceived as risky, which, in turn, leads to a lower potential level of take-up. Likewise, a recent EU survey on the use of technologies based on AI concluded that 33 % of enterprises find liability for potential damages to be one of the major external challenges to AI adoption in the EU.

Background - concerns over existing framework (PLD and traditional liability regimes)

The European Commission has identified several challenges in applying the Product Liability Directive (PLD) and traditional liability regimes to digital content, software, and data. These challenges include:

- ▶ 1. Unclear Classification: It's uncertain whether intangible elements like digital content, software, and data can be classified as products under the PLD. This leads to legal ambiguity in compensating injured parties for damage caused by software, including updates, and determining liability.
- ▶ 2. New Technological Risks: Emerging technologies introduce novel risks, such as vulnerabilities to data inputs that affect safety and cybersecurity. However, the PLD only covers compensation for physical or material damage, leaving a gap in addressing these new risks.
- ▶ 3. AI Complexity: AI systems possess unique characteristics, such as opacity, lack of transparency, autonomous behavior, continuous adaptation, and limited predictability. These traits make it extremely challenging for victims to meet the burden of proof required for a successful claim under current liability rules. Victims usually need to demonstrate damage, fault, and a causal link, but with AI, identifying fault, defects, and causality is exceptionally complex, making it difficult to obtain compensation.

reform

ongoing reform of the EU liability framework that applies to AI. The reform is twofold. On one hand, it consists of the reform of the product liability directive and is presented in the PLD Proposal. On the other hand, at the same time, the Commission has published a proposal for AI Liability Directive

What is the AI Liability Proposal

- ▶ **Directive-** opens the path to harmonization and sets regulations for the entire EU while giving Member States latitude in how they implement the directives. Procedural rules not material
- ▶ This is especially important when it comes to the proposal for an AI system liability regime-is predicated on ideas like damage and responsibility, which vary substantially amongst legal systems meaning that the result of applying the AI liability Directive can be different from one member state to another not providing accomplishment of the idea for harmonization.
- ▶ AILD Proposal seeks to harmonize procedural questions, such as disclosure of evidence and burden of proof across Member States' national liability regimes for the purposes of AI liability, while largely tying these instruments to violations of the AI Act.

Key objectives



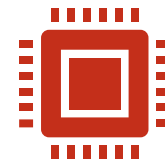
Lay down uniform requirements for non-contractual civil liability for damage caused by the involvement of AI systems.



promote the rollout of trustworthy AI, to harvest its full benefits for the internal market by ensuring victims of damage caused by AI obtain equivalent protection to victims of damage caused by products in general



reduce legal uncertainty for businesses developing or using AI regarding their possible liability exposure and prevent the emergence of fragmented AI-specific adaptations of national civil liability rules.



to complement and modernize the EU liability framework to introduce new rules specific to damages caused by AI systems. The new rules intend to ensure that persons harmed by AI systems enjoy the same level of protection as persons harmed by other technologies in the EU.

Scope of the proposal



Harmonize non-contractual civil liability rules for damage caused by artificial intelligence (AI) systems (Article 1).



The AI liability directive would not define AI, but refer to the same general concept of AI as in the AI act and particularly its definition of 'AI systems'. The new rules would apply to damage caused by AI systems, **irrespective of whether they are defined as high-risk or not** under the AI act



The AI liability directive would not affect existing rules laid down in other EU legislation, particularly the EU rules regulating conditions of liability in the field of transport, the proposed revision of the Product Liability Directive or the Digital Services Act.



while the AI liability directive does not apply for criminal liability, it may be applicable for state liability given that state authorities are subject of the obligations in the AI act

Key provisions- Presumption of causality

- ▶ Article 4 lays down a rebuttable presumption of causality establishing a causal link between non-compliance with a duty of care under Union or national law (i.e. the fault) and the output produced by the AI system or the failure of the AI system to produce an output that gave rise to the relevant damage.
- ▶ Such presumption of causality would apply when the cumulative following conditions are met:
 - ▶ 1. non-compliance with a certain EU or national obligation relevant to the harm of an AI system caused the damage (Article 4 (1)(a)).
 - ▶ 2. it must be reasonably likely that, based on the circumstances of each case, the defendant's negligent conduct has influenced the output produced by the AI system or the AI system's inability to produce an output that gave rise to the relevant damage (Article 1(b)).
 - ▶ 3. it must be reasonably likely that, based on the circumstances of each case, the defendant's negligent conduct has influenced the output produced by the AI system or the AI system's inability to produce an output that gave rise to the relevant damage (Article 1(b)).



Reasonably likely

The academics stress that the question whether the requirement of 'reasonably likely' is met would rest on a subjective assessment by national judges on a case-by-case basis. This may affect legal certainty and cause fragmentation across the EU depending on national tort law traditions.

When the presumption will apply - Types of risk in AI systems

providers and users of high-risk AI systems

- ▶ In the case of a claim for damages against a provider of a high-risk AI system, national courts must therefore presume the causal link between the non-compliance of these requirements and the output produced by the AI system, or the failure of the AI system to produce an output that gave rise to a relevant damage

not a high-risk AI system

- ▶ the presumption would only apply where the national court considers it excessively difficult for the claimant to prove the causal link (Article 4 (5)). Furthermore, the presumption of causality does not apply if the defendant proves that the claimant has sufficient evidence and expertise to prove the causal link between the fault of the defendant and the output produced by the AI system or the failure of the AI system to produce an output that gave rise to a relevant damage (Article 4 (4)).

Strict liability?

- ▶ Part of the academics recommend introducing a strict liability regime for high-risk AI systems and a complete reversal of the burden of proof for other AI systems to clarify the scope and application and offer greater legal certainty.
- ▶ The draft text does not make it easy for injured parties to establish a presumption of causality, given the heavy burden of proof they face, and argues that amendments should be introduced to facilitate the mechanism of redress available to victims of AI harm.
- ▶ For the theoreticians it is questionable whether a fault-based liability regime would succeed in simplifying victims' claims even with a regime of presumption of causality. They argue that AI systems can be so complex that even when a user complies with their duty of care, damage can still arise and it is not clear who will be held liable for such damages and on what grounds

Key provisions - Disclosure of evidence

- ▶ the AI liability directive would give national courts the power to **order disclosure of evidence about high-risk AI systems** that are suspected of having caused damage (Article 3(1)). The new rules would help victims to access relevant evidence to identify the person that could be held liable, for instance, when damage is caused because an operator of drones delivering packages does not respect the instructions for use or because a provider does not follow requirements when using AI-enabled recruitment services.²³ Accordingly, companies responsible for high-risk AI systems would be required to disclose a range of information including specific documentation, information and logging requirements

Implications for developers

- ▶ The Commission discards such a reversal of the burden of proof to avoid exposing providers, operators and users of AI systems to higher liability risks, which could hamper innovation in AI-enabled products and services. Under the proposed approach, the victim would, instead, still bear the burden of proof, but the presumption of causality would result in a targeted alleviation of the burden of proof regarding the question as to how or why an AI system reached a certain harmful output. This approach would relieve victims of the need to demonstrate the inner workings of the AI system at stake.

Relations with national law

- ▶ The directive would lay down EU rules for presumption of causality but not harmonise rules regarding which party has the burden of proof or which degree of certainty is required as regards the standard of proof. This remains a Member State competence within national laws. Furthermore, the proposed directive follows a **minimum harmonisation approach**. This would allow claimants to invoke more favourable rules under national law (e.g. reversals of the burden of proof under national fault-based regimes or national no-fault liability), for instance in cases of damage caused by AI systems
- ▶ Member States may adopt or maintain national rules that are more favourable for claimants to substantiate a non-contractual civil law claim for damages caused by an AI system, provided such rules are compatible with Union law

Challenges and concerns

- ▶ Need for strict liability?
- ▶ it is questionable whether a fault-based liability regime would succeed in simplifying victims' claims even with a regime of presumption of causality
- ▶ the proposed directive follows a **minimum harmonisation approach**. This would allow claimants to invoke more favourable rules under national law (e.g. reversals of the burden of proof under national fault-based regimes or national no-fault liability), for instance in cases of damage caused by AI systems
- ▶ AI Complexity