



ELB Blogpost 48/2022, 6 November 2022

Tags: Proposed AI Act (COM(2021) 206 final); Proposed Directive on adapting non-contractual liability rules to artificial intelligence (COM(2022) 496 final); Proposed Directive on liability for defective products (COM(2022) 495 final)
Topics: Data protection and digital governance

The Proposed EU AI Liability Rules: Ease or Burden?

By Samar Abbas Nawaz

“How should anyone be held liable for a harm caused by Artificial Intelligence (AI) systems?” is an oft-raised concern around the use of AI systems. After proposing a regulation for AI systems in 2021, the European Commission has now addressed this critical issue of liability. Regulating AI liability is not an easy task, and the recent legislative efforts of the Commission make this even clearer. Ensuring that AI systems are safe to use is crucial for law makers. On the basis of two proposed directives, the Commission aims to protect victims of AI harm by setting some rules for imposing liability. It is thus imperative to examine to what extent these rules address the concern of liability and make it easier for injured parties to receive damages for the harm incurred.

Background

In 2020, the Commission published a [White Paper on AI](#) which set out a commitment to address various risks linked to the use of AI systems. This was followed up in 2021 by a [proposal](#) for an AI Act (AIA) which is yet to be adopted. The proposed regulation sets out important requirements to minimize harmful risks to individuals by differentiating between prohibited and high-risk AI systems. AI systems that are prohibited include real-time biometric identification, social scoring, or manipulative AI systems (Article 5). Stricter requirements apply to high-risk AI systems. A ‘high risk’ AI system constitutes any AI-powered system or safety component of a system which is bound by a conformity-assessment requirement under harmonized EU regulations (Article 6) These systems may

only be used if they comply with prescribed risk management system, transparency, human oversight, appropriate data quality, and effective cybersecurity requirements (see Chapter 2). The proposed AIA thus operates to ensure the safety of AI systems against possible risks.

However, what happens once those risks materialize and someone is harmed? How can anyone be held liable? To deal with these questions, the Commission recently proposed two directives: namely, an [AI Liability Directive](#) (AILD or Directive on adapting non-contractual civil liability rules to artificial intelligence) and a [revised Directive on liability for defective products](#) (PLD). The AILD paves the way for proving someone's fault when an AI system causes harm. In legal parlance, it can be seen as a mechanism for 'fault-based liability'. The revised PLD is useful in instances where harm is caused by a defect in an AI system, not by virtue of someone's fault. Hence, it establishes 'no-fault liability' rules. The idea behind this endeavour is to enhance public trust in AI technology and provide legal clarity to businesses involved in developing these systems (see Sections 1.1. of the Commission's explanatory memoranda to the PLD and AILD).

AI Liability Directive (AILD)

Major aspects that the proposed AILD covers are the presumptions and provision of information about the AI system to an injured person. Two presumptions can be found in the proposed Directive. One extends to non-compliance with the duty of care and another for the nexus between fault and the output, termed as 'presumption of causality' (Articles 3 and 4). Simply put, the presumption of causality infers that the harm caused by an AI system was indeed due to the fault of the provider or user. This presumption activates if the provider or user, whoever the defendant in the case may be, fails to provide information about the relevant system even after a national court has ordered such a disclosure (Article 3(5)). A 'provider', in the context of the AILD framework, is defined as a manufacturer or any other person who places the given AI system on the market under its own trade name (Article 2(3)). A 'user' can be anyone using the AI system under their authority. (Article 2(4))

However, the defendant may rebut the presumption by demonstrating that sufficient information regarding the AI system could be accessed to prove the fault, if any, on its part (Article 4(4)). Another condition for the successful imposition of liability on the provider is that the national court must infer that the damage caused by the given AI system was 'likely' caused by its fault (Article 4(1)(b)). The provider would be held liable if they violate their duty of care under relevant law and any of the requirements of the AIA relating to the quality of data, transparency, human oversight, cybersecurity, or lack of withdrawal of a given system when the fault was identified (Article 4(2) AILD). On the other hand, a user would become liable if they violated the instruction of use or if they exposed

the system to irrelevant input data (Article 3). However, the claimant would need to prove such likelihood to avail the benefit of this presumption.

Regarding the provision of information about the harmful AI system, the AILD allows an injured person to seek information through the national court (Article 3). A similar provision on the disclosure of information can also be found in Article 13 of the proposed AIA to ensure transparency of AI systems. A major difference between these requirements is that the AIA requirement is directed towards the user of the system, whereas the AILD requires the disclosure of information to any victim of AI harm. The AILD thus aims to make it easier for claimants to prove fault based on the disclosed information, and thereby to seek compensation for any damage incurred. For instance, if an AI-powered drone in civilian settings malfunctions and injures a bystander, then such a bystander could claim damages based on the fault of a natural person. The AILD would allow the claimant to seek information about the drone from its provider or user to identify the fault.

Revised Product Liability Directive (PLD)

A revision of the PLD has been long overdue, considering that the original Product Liability Directive was adopted in 1985 ([Directive 85/374/EEC](#)) and thus had several shortcomings when applied to AI systems. The proposed Directive shall replace it. A significant aspect of this revision is that it brings software within its scope of application. This is [seen as a deviation](#) from the traditional position that only manufacturers or importers of a product can be held liable and not software developers. In contrast to the AILD, the revised PLD allows holding the provider liable on the basis of a defect in the product instead of on the basis of fault. Therefore, if an AI-powered drone causes harm to any individual because of some defect then the claimant could seek remedy under the revised PLD.

Article 9 PLD recognizes two presumptions: a presumption of defectiveness (Article 9(2)), and a presumption of a causal link between the defectiveness and the damage (Article 9(3)). The presumption of defectiveness is triggered in three instances (see Article 9(2)), if:

- (1) the provider fails to disclose system information; or
- (2) the claimant shows an obvious malfunction; or
- (3) the claimant shows a violation of safety rules on the provider's part.

The presumption of a causal link activates if the damage is 'typically consistent' with the defect. Both presumptions can be made simultaneously if the technical complexity of the system renders it difficult to identify the defectiveness (Article 9(4)). However, to be able to rely on these presumptions, the claimant would still have to prove the likelihood of the defectiveness and that the product contributed to damage. However, the presumptions may be rebutted by the provider by showing that such excessive technical difficulties do not exist (Recital 34).

Challenges

One of the criticisms facing especially the proposed AILD is that, despite the provision of relevant information, it can still be very difficult to prove fault in complex systems. Especially, as some AI systems behave autonomously and share complexity in their functions where the reason of any output cannot be understood easily. This challenge speaks to the 'black box' algorithm issue which arises when the intricacies of an AI system render it difficult to understand the input which leads to a certain output. In these circumstances, what good would information about the system do to an injured party?

While the proposed AILD mentions autonomy as a problem in understanding the system—in Recitals 3, 27, and 28—it does very little to make it easier for injured parties to establish a presumption of causality. An injured party will still face a heavy burden of proof under the AILD: from providing evidence in support of the plausibility of the claim (Article 3(1)) to identifying non-compliance with the AIA requirements (Article 4(1)(a)), and demonstrating a link between the action of the AI system and the damage sustained (Article 4(1)(c)). It might be also quite arduous to prove non-compliance with the requirements set out in the proposed AIA. For instance, it may not be easy to prove that the datasets used in the development of an AI system, or the accuracy levels of given system are indeed inadequate. Hence, the proposed AILD, at best, provides very limited procedural ease for injured parties. More needs to be done to facilitate effectively the mechanism of redress available to victims of AI harm.

One way to deal with this issue is to use the defect-based remedy under the revised PLD where no fault needs to be proven. However, under the PLD, compensation can be sought only in case of material harm (Article 4(6)). It means that AI systems involved in, for instance, credit-scoring by any financial institution which could harm individuals in a non-physical way could not be challenged on the basis of being defective. In this case, an aggrieved person would have to prove fault through the AILD to receive compensation. Ursula Pahl, the Deputy Director of the European Consumer Organization (BEUC) has already [voiced this concern](#) whilst talking about the proposed directives. In this context, the Commission's approach is more beneficial for developers, some of whom have [opposed](#) strict liability against immaterial damage during the preparation of these proposals. In addition to this hinderance, the PLD also requires claimants to prove the likelihood that harm was done due to the defect of the system to avail the presumption of defectiveness (Article 9). This raises a question about the parameter of such likelihood that a claimant must meet, especially when a given system is too complex to understand.

Conclusion

The Commission's proposed AI liability directives primarily make it easier to seek information about AI systems to establish liability. This, in a way, hardwires transparency requirement for AI systems which is also contained in the proposed AIA. However, while doing so, the proposed liability directives also burden claimants with some tricky obstacles to overcome; be it having to establish the fault or presumptions of defectiveness and causality, and the nexus between the harm and defect or fault, which can be found in the AILD and PLD. The Commission's proposals are at the initial legislative stage and will likely go through various modifications before their final adoption. Moreover, potential modifications in the proposed AIA might also still change the implementation of the proposed liability directives since the directives rely on the definitions and requirements set out in the AIA. The legislative efforts towards establishing AI liability are an important step in the right direction to regulate AI effectively. However, it remains imperative to be cautious of the complexities involved in AI systems especially when the goal is to protect injured parties.