EU AI Liability Directive

Strengths, Weaknesses, and Suggestions for Improvement

(you can tell this presentation is about law based on how soulless it looks)

What is the EU AI Liability Directive?

The EU's new 'AI Liability Directive' proposes a ***targeted reform of national fault-based liability regimes*** with respect to damages caused by AI systems, with the aim of enhancing consumer trust in AI and ensuring successful innovations across the EU.

Unlike hard laws and regulations (which are directly applicable to all Member States immediately after their introduction/implementation), directives are ***not directly and immediately applicable to Member States***: they have to first be transposed into the (potentially varying) national laws of particular jurisdictions.

Accordingly, this proposal lays out general EU rules for presumption of causality, but ***does <u>not</u> 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 to be adapted according to national laws.

Article 3: Disclosure of Evidence - What does it say?

1. National courts are empowered to disclose relevant evidence at its disposal about a specific **high-risk AI system** that is suspected of having caused damage.

4. National Courts shall limit the disclosure of evidence to that which is **necessary and proportionate** to support a potential claim or a claim for damages. To determine whether evidence is necessary and proportionate, it will consider the evidence of all parties, including third parties concerned.

5. Where a defendant fails to comply with a court order to disclose evidence, **the court shall presume the defendant's non-compliance with a relevant duty of care.** The defendant shall have the right to rebut that presumption.

Strengths and Weaknesses

Strengths:

- ***For Claimants*:** If the burden of proof was to stay on the claimant in the situation where the defendant cannot provide basic evidence *(necessary and proportionate)*, the claimant may not even know how the system works, meaning it would be practically impossible to provide evidence of their own. The Directive correctly ***reverses the burden of proof by requiring a presumption of non-compliance*** in this situation to help claimants access the relevant evidence to prove their claims.
- ***For Defendants*:** The ***consideration of all relevant interests*** - companies, developers, the public at large - prevents tension with other legal regimes, such as intellectual property law and national security

Weakness: The court can only requisition evidence for claims relating to **high-risk** systems: defined in **Annex III** of the EU AI Act, qualified with **Article 6(2)** - they are **high-risk** only if they pose a ***significant risk* of harm to: i) the health, safety or fundamental rights of natural persons, or, ii) for AI systems under Annex III point 2 (critical infrastructure), the environment**.

The Directive should consider potential emergent properties of AI which may currently be “limited risk”, and the fact that limited risk AI may still cause tortious harm.

Article 4: Presumption of Causation Test

Three stage test (must prove all three steps and can be rebutted by the defendant)

1. Defendant has not complied with a duty of care that would protect against the damage occurred
2. It is reasonably likely that the defendant's fault influenced the output produced by the AI
3. The claimant has shown that the AIs output gave rise to the harm

Weaknesses and Strengths

Strengths

- For Claimants: ***Reduced burden of proof*** with respect to causation will make it easier for people alleging injury from AI to succeed in bringing claims (claimants do not need to demonstrate the inner workings of the AI system)
- For Defendants: ***Does not entail a complete reversal of the burden of proof,*** according to which the victim no longer bears *any* burden of proof and it is for the person liable (i.e. the defendant) to prove that the conditions of liability are not fulfilled, thereby avoiding the exposure of AI providers to unnecessarily high liability risks, which could hamper innovation in AI-enabled products and services.

Weaknesses

- Using traditional concepts of law that do not reflect the special characteristics of AI models: especially emerging capabilities and the ‘black box’ phenomenon .
- Proving “reasonably likely” is extremely hard in the AI context from the perspective that many cases will arise where the claimant cannot prove causality because of the emerging capabilities of AI and the lack of foreseeability of these capabilities.
- For these reasons, high-risk “illegitimate harm” models should be subject to ***strict liability (i.e. no-fault liability)***

Suggestions for improvement

Disclosure of Evidence for Claims: Courts must be able to requisition evidence from providers of non-high-risk AI - change Article 3(1).

Liability Attribution: Attribution Models: these models may include concepts like joint liability or proportional liability based on the level of control and foreseeability. In case of low control over outputs, such as for highly general and capable systems, we suggest **Legal causation vs. Factual causation** should be the default approach in a claim dispute

Hacker‘s “illegitimate harm” and “legitimate harm” AI Model Distinction:

Change Article 4 to include strict liability for “illegitimate harm” AI models (i.e. autonomous cars, medical AI) and “legitimate harm” AI models (i.e. Credit scoring AI, employment hiring AI) stay with the current presumption of causation test.

Insurance and Financial Responsibility:

- Compulsory Insurance Schemes: Establishing mandatory insurance for AI operations to cover potential damages, similar to automobile insurance.
- Victim Compensation Funds: this is especially in situations where specific liability is difficult to ascertain

Thank you for your attention!