

The EU AI Liability Directive Will Change Artificial Intelligence Legal Risks

October 24, 2022

On September 28, 2022, the European Commission released a proposal to change the legal landscape for companies developing and implementing artificial intelligence in European Union (EU) Member States. This [AI Liability Directive](#) would require Member States to implement rules that would significantly lower evidentiary hurdles for victims injured by AI-related products or services to bring civil liability claims. Most importantly, the Directive would create a “presumption of causality” against the AI system’s developer, provider or user.

The proposed AI Liability Directive should be seen as part of a broader package of EU legal reforms aimed at regulating AI and other emerging technologies. The other parts of that package include the [draft EU AI Act](#), which aspires to establish the world’s first comprehensive regulatory scheme for artificial intelligence, and the [Digital Services Act](#) (“DSA”), which is set to transform the regulation of online intermediaries.

In this Debevoise Data Blog post, we explore the key elements of the proposed AI Liability Directive, as well as steps that businesses should consider to enhance their AI governance and compliance programs in anticipation of these changes.

Civil Liability and the “Closed Box” of AI

The AI Liability Directive is meant to (a) modernize the EU’s existing liability regime as applied to AI, (b) reduce uncertainties for businesses operating across jurisdictions and (c) instill confidence in consumers interacting with emerging technologies. It would apply broadly, covering any claims relating to AI products or services involving a “fault or omission of any person (providers, developers, users).” The Directive does not create new legal obligations, but it could impact many types of claims covered by national law in the EU, including harms to life, health or property, privacy, equality and non-discrimination, by altering the evidentiary requirements in fault-based liability regimes. By contrast, strict liability claims that only require the victim to show that a harm resulted from the use of a product are covered by updates to the EU’s [Product Liability Directive](#), which were also proposed on September 28, 2022.

Claims for harm caused by AI systems can be complex and opaque. Claimants often struggle to prove causality because they have minimal visibility into how that AI system actually works—what is often called the “black box” or “closed box” effect. As a result, they may be unable to show that a specific defect, design flaw or other malfunction in the system was *actually the cause* of their injury.

For example, if a self-driving car collides with a pedestrian, the claimant knows that she has been injured by the collision. But she may not be able to prove what caused it without an enormous amount of time and effort, if at all. Was it a malfunction in the car’s pedestrian-sensing algorithm, a software bug resulting from the car owner’s failure to download an update, a hack of the car’s systems or an error on the part of the driver? As the Commission notes, victims of AI-related harms may “incur very high up-front costs and face significantly longer legal proceedings, compared to cases not involving AI,” which might deter them from pursuing claims altogether.

Key Provisions of the AI Liability Directive

The AI Liability Directive would require EU Member States to update their national civil liability regimes to provide consumers with the “same standards of protection when harmed by AI systems as they would be if harmed under any other circumstances.” Specifically, the Directive would require States to implement two core changes to their national laws:

Easing the Burden of Proof through a “Rebuttable Presumption of Causality”

By reducing the burden of proof, the Directive would make it easier for people alleging injury from AI to succeed in bringing claims. It creates a “rebuttable presumption of causality,” under which a claimant would not have to prove that the defendant was at fault so long as: (a) the AI system’s output (or failure to produce an output) was reasonably likely to have caused the damage; (b) that damage or harm was caused by some human conduct influencing the AI system’s output, and (c) the conduct did not meet a duty of care under EU or national law that was directly intended to protect against the damage that occurred. Notably, the failure to meet this duty of care can be established if the defendant did not comply with regulations applicable to the AI, which would include obligations for “High-Risk AI Systems” as defined under the EU AI Act once it is in effect. If the presumption is invoked, the defendant can rebut it by producing evidence showing its fault could not have caused the damage.

For the self-driving car collision, this means that the pedestrian would not need to clear the steep hurdle of identifying what specific aspect of the self-driving car caused her injury, nor prove who, exactly, is responsible for the collision. Rather, she would only

need to show that (a) the car's manufacturer or the developer of the car's AI component(s) failed to comply with a certain legal duty (e.g., failing to audit the car's computer vision algorithm to determine how well it performs), (b) it is "reasonably likely" there is a causal link between that failure and the AI system's performance (e.g., because the computer vision may not have worked), and (c) the AI system's failure "caused" the accident (e.g., by failing to trigger a brake decision at the right moment). With that showing established, the burden would shift to the car manufacturer to rebut the presumption of causality by demonstrating that its AI component performed as intended (e.g., by triggering the correct brake decision, which was then overridden by the human driver).

Increasing Transparency through Disclosure of AI-Related Evidence

The Directive would also make it easier for potential claimants to obtain court orders that mandate disclosure of relevant evidence concerning High-Risk AI Systems, which those claimants could use to identify the actors who are potentially liable for these harms. By focusing specifically on High-Risk AI Systems, the European Commission implicitly underscores the importance of implementing the documentation, logging, record-keeping and transparency obligations [contemplated by the AI Act](#). So, if a company's High-Risk AI System results in a fault-based liability claim, the company could be required to disclose information ranging from audit logs to risk assessments to the purported victim. And critically, if the company fails to disclose this evidence (including because it never took steps to document or preserve it), the Directive would also allow a court to invoke a presumption that the defendant had not complied with its duty of care under EU or national law.

Timeline for Adoption and Enforcement

The next step for the Directive is for the European Parliament and Council to consider and adopt the draft text. Once that is finalized into the final Directive, it would enter into force soon thereafter. Member States would then have two years to implement this new liability framework under domestic law. The Directive would not apply retroactively, but rather, only to claims that arise after this two-year period elapses.

Key Steps Companies Can Take to Prepare for European AI Liability Claims

The details of the AI Liability Directive are subject to change as EU lawmakers consider the draft text over the coming months. But companies investing heavily in AI may want to start preparing for this new civil liability framework, which could expose them to

significant risk of civil claims and reputational damage. Steps that such companies may consider include:

- **Preparing for Compliance with the EU AI Act:** The European Commission has described the proposed Directive and the draft AI Act as “[two sides of the same coin](#),” noting that the AI Act seeks to compel companies to develop and use AI in ways that will avoid causing harm, while the Directive applies to situations in which AI has already caused harm. Accordingly, to mitigate future civil litigation risk, companies could evaluate their AI Act-compliance strategy by taking an inventory of their AI models to make sure they know which, if any, would qualify as High-Risk AI Systems under the Act and then assessing what needs to be done to ensure compliance with the Act’s requirements.
- **Improving Documentation and Audit Capabilities:** Because the presumption of causality is rebuttable, companies may be able to show that a different cause is actually responsible for the claimant’s injury. Maintaining robust documentation of model testing and activity logs of model performance may position companies to better defend themselves from claims against their AI systems.
- **Training:** To reduce the risk of litigation, companies should consider providing training for individuals involved designing, developing, implementing and operating high-risk AI to detect and mitigate potential problems.
- **Incident Response Planning and Testing:** The Directive specifically notes that providers or users of High-Risk AI Systems may face a presumption of causality if they fail to appropriately ensure the security of their AI systems, monitor the AI system while in use or suspend or interrupt the use of the AI system in the event of a significant risk or incident (as required under the AI Act). In light of this, companies should consider creating a plan for swiftly identifying and escalating AI incidents and responding to allegations of harm caused by their AI, as well as testing that plan through an AI tabletop exercise.

To subscribe to our Data Blog, please click [here](#).

The [Debevoise Artificial Intelligence Regulatory Tracker](#) (DART) is now available for clients to help them quickly assess and comply with their current and anticipated AI-related legal obligations, including municipal, state, federal and international requirements.

* * *

Please do not hesitate to contact us with any questions.



Avi Gesser
agesser@debevoise.com



Robert Maddox
rmaddox@debevoise.com



Anna R. Gressel
argressel@debevoise.com



Tristan Lockwood
tlockwood@debevoise.com



Frank Colleluori
facolleluori@debevoise.com



Michael Pizzi
mpizzi@debevoise.com