Thou Shalt Pay

All you need to know about AI liability in 2026 | Edition #261

LUIZA JAROVSKY, PHD

JAN 05, 2026 · PAID

20 3 Share

Marinus van Reymerswaele, “The Moneychanger and His Wife,” 1539 (oil on panel, modified)

👋 Hi everyone, **Luiza Jarovsky, PhD**, here. Welcome to the **261st** edition of my newsletter, trusted by more than **88,600** subscribers worldwide.

Welcome to our first edition of 2026. Some are saying this is the year AI will transition from hype to pragmatism. I sincerely hope this will be the year when enforcement and oversight finally catch up.

Thou Shalt Pay

Many in AI seem to have a hard time understanding legal liability and the fact that companies often have to pay for the harm they cause.

Every now and then, I read baseless comments that sound more like wishful thinking from people who have absolutely no idea how liability actually works and who think

that, for some reason, AI should be immune from legal oversight.

So let me break down where we currently are and what we should aim for in terms of holding AI companies accountable for the harm their products cause.

-

First, regardless of how shiny and 'AGI-like' an AI system might be, it is, from a legal perspective, a product (and sometimes also a service), just as a car, a computer, and a toothbrush are.

There are rules applied to all sorts of products, such as the types of materials or manufacturing techniques allowed, the standards to be followed, the required design, and their intended performance.

There are also product liability rules, based on the general product rules, that specify what will happen when a product causes harm.

Product liability addresses questions such as **when a company must compensate victims** or their families if its products cause harm.

These rules will differ depending on the product and the state or country. They are usually proportional to the known or expected risks posed by the product.

In general, there are two main liability systems: strict liability and fault-based liability.

Strict liability means that even if there was no negligence or intent on the part of the company, it may still be held responsible for the harm its product causes (provided that other elements are proven, usually the 'defect' and the causal relationship between the harm and the defect).

In the **fault-based liability** system, on the other hand, negligence, recklessness, or intent by the company must be proven by the victim or the victim's legal representative.

Now, here is something that seems to blow the minds of many in AI:

In both liability systems, there may be cases in which a person was careless in using a product or used it in a risky or 'dumb' way, and the company that manufactured it will still be held responsible for the harm its product caused.

Yes, you read it correctly.

Even if the person used the product inadequately, the company behind it might have broken product rules (such as mandatory warnings, product design, or safety measures), which would still make it responsible.

This is true especially with new technologies, where most people do not know how to use them properly or navigate the risks. (I would say even more when general-purpose AI systems such as ChatGPT are involved, as they are designed for many possible uses).

In these cases, misuse is often expected, and there will be greater scrutiny of warnings, guardrails, built-in safety, and design requirements.

The details will depend on the product's regulatory requirements and the product liability rules of the specific jurisdiction, but, in general, that is how product liability works.

In AI, it means that even if a person uses an AI system in an inadequate, risky, or dumb way, the AI company might still be held responsible for the harm the AI system caused.

So, what are the accepted liability rules for AI systems today? If a person takes their own life after being encouraged by an AI chatbot to do so, does the AI company have to compensate the family?

There is still a lot of uncertainty and many gray areas around AI liability, especially in areas such as general-purpose AI chatbots, given that the technology only became widely available (and started being used by hundreds of millions of people) about three years ago.

There are, however, a few high-stakes lawsuits currently underway against OpenAI and CharacterAI involving AI chatbot-related suicides and mental health harm that will likely offer answers and help set clearer boundaries in the field of AI liability, hopefully as early as 2026.

Among the lawsuits against OpenAI are those filed by the Raine family, the Social Media Victims Law Center, and, more recently, by the estate of Suzanne Adams.

Among the lawsuits against CharacterAI are those filed by the Social Media Victims Law Center and Megan Garcia.

Even though most of the lawsuits above include additional claims (beyond liability), their main focus is liability and holding AI companies accountable for the harm their products caused.

Regarding the outcome of these lawsuits, in my opinion, the insufficient guardrails (that allowed shocking instances of suicide encouragement to happen, such as in the Raine case) will likely be enough to establish the defectiveness of the AI chatbots for purposes of legal liability.

The core of the legal dispute will revolve around whether the AI chatbot endorsement and encouragement in each case counts as the cause or a legally sufficient causal contribution in cases where mental health issues and suicide are involved.

-

How can we solve the liability puzzle for cases involving AI chatbots and beyond?

An important bottleneck and obstacle to clearer liability rules for AI is the lack of effective regulation of AI in general.

General-purpose AI chatbots, for example, are barely regulated. Without legal frameworks governing them, we cannot answer questions such as:

- What are the minimum safeguards that should be available by default to every user of the product?
- What type of guardrails should be activated when minors are using it?
- What should happen when people engage in potentially dangerous, emotionally charged, or sensitive conversations? Should there be barriers or limits?
- What should the interface design of an AI chatbot look like?
- What type of transparency warnings should be mandatory?
- Should there be use cases or interactions that are blocked by default?

In most places, these rules do not exist (that is why I often say we still live in the regulatory Wild West of AI), and companies usually do the bare minimum.

Without these basic product safety rules, liability will remain a gray area, and when people are harmed or commit suicide after interacting with an AI chatbot, the victim or their family will have to go to court to demand justice.

Digging deeper, to have better product safety laws and standards for AI, we desperately need more data: cognitive, behavioral, psychological, neurological, and from various other fields, measuring, mapping, tracing, and dissecting human-machine interactions.

We need to know much more about how people are affected by interactions with AI systems, how they may be harmed, and what the real risks are.

It is not possible to draft effective rules if we do not know what the risks are.

-

Lastly, AI laws, both product safety and liability laws, must be drafted carefully.

As I have written a few times in this newsletter, because the field of AI is evolving extremely rapidly, future-proofing techniques are essential to ensure that the law remains relevant for at least a few years after it is enacted.

The EU AI Act, for example, did not adequately assess the risks posed by certain types of AI, including general-purpose AI systems.

In addition to that, its future-proofing mechanisms are scarce and much less comprehensive than those of the Japanese AI law, for example.

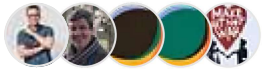
Due to these (and various other factors that I discuss in depth in my AI Governance Training), the EU ended up with a legal framework that is much weaker than it could have been. And the proposed Digital Omnibus attempts to further weaken it.

-

Regardless of the outcome of the lawsuits I mentioned above, AI liability will remain a gray and uncertain area until we enact strong, comprehensive, and dynamic product safety rules and standards for AI.

And besides granular data, an essential element for productive lawmaking in any field is sufficient political will.

Unfortunately, as of now, there does not seem to be an appetite for effective AI regulation in most places in the world. I hope this will change.



20 Likes · 3 Restacks

← Previous Next →