Policy Report

AI Liability Along the Value Chain

BEATRIZ BOTERO ARCILA

SUPPORTED BY MOZILLA

Mozilla

Policy Report: AI Liability Along the Value Chain

Policy Report

AI Liability Along the Value Chain

BEATRIZ BOTERO ARCILA

Table of Contents

Introduction and summary

Policymakers around the world are increasingly preoccupied with identifying mechanisms to better assign accountability and liability throughout the AI value chain. Particularly in the EU, discussions around civil liability and AI received significant attention after the proposal of an AI Liability Directive (AILD) in 2022. While this proposal was recently withdrawn by the European Commission, the challenges posed by AI for civil liability and harmed individuals' ability to seek redress remain more relevant than ever amid increasing adoption of AI across sectors. This report thus seeks to provide more conceptual clarity to these challenges and provide recommendations on what an effective AI liability framework could look like.

Though it is common to think of AI systems as a singular tool, AI systems are often developed and deployed in a value chain that involves numerous actors that participate throughout the stages of creation, fine-tuning, and implementation of these technologies, or that sell and supply key components such as pre-labeled data.

When designing a liability system for this type of multi-party scenario, there are many questions to consider: should all parties in the value chain be held equally liable when harm occurs? Or should each actor only be held liable for the extent to which they are responsible? How easy is it to establish the contribution of each party? (Spoiler alert, it may be very hard.) Another question lawyers will be familiar with is what is the right standard — should AI actors be held liable only when they fail to take the right safety measures? Or should they be held liable regardless of whether they took safety measures, simply because by developing or deploying an AI system or model they created a risk?

This Report discusses these questions and the complexities of assigning liability along the AI value chain, given the involvement of multiple actors in the design, development, and deployment of AI systems. The Report explores various configurations of AI value chains, the roles of different actors, and how companies allocate liability amongst them via contracts and terms. It then examines different policy choices for designing liability regimes.

The Report argues that given the particularities of the AI value chain, the most desirable baseline regime is one where all parties in the AI value chain are equally responsible for harm when they act with neglect or fail to take the appropriate harm-mitigation measures. This proposed framework creates the best incentives for all the actors in the

value chain to take adequate safety measures.[1] This general regime, however, still faces important challenges and therefore there may be important exceptions where AI actors involved in the deployment of particularly dangerous AI systems are held liable under a strict liability standard.

The framework proposed here resembles to a certain degree the AI liability that was discussed in the context of negotiations around the EU's AILD. This Report thus ends by discussing that framework, and the limitations and lessons it may provide for Europe itself and other jurisdictions.

The Report is structured as follows:

- **Section 1: Artificial intelligence accountability and liability today**
  Introduces the issue of AI accountability for liability law, and how it is enhanced by AI opacity and complexity.

- **Section 2: The problem of many hands in AI accidents and different kinds of AI value chains**
  Expands on the "problem of many hands" and why it represents a particular challenge for AI liability. It also presents different configurations of AI value chains and the actors that participate in them.

- **Section 3: The structure of roles and responsibilities along AI value chains**
  Discusses how actors in the AI value chain already establish obligations and allocate liability amongst them, for example, through the terms of use and use licenses, but also through their design choices and commercial relations.[2]

- **Section 4: Liability rules as applied to the AI value chain**
  Discusses different policy choices to consider when designing liability laws and systems for accidents that involve several actors and offers a series of criteria for allocating liability in the AI value chain. The main policy options to consider are three: (1) What is the liability regime? (2) Joint and several or several liability? (3) When and how to address information asymmetries?

- **Section 5: The EU liability regime**
  Discusses the proposed AI liability regime in the European Union using the frame-

1 A similar conclusión has been reached by Miriam Buiten, Alexandre de Streel, Martin Peitz "The law and economics of AI liability," *Computer Law & Security Review*, Volume 48, April 2023.

2 This framework, which holds that responsibility and liability are structured and governed by architecture, the market, law and norms, echoes Larry Lessig's pathetic dot theory. Larry Lessig, Code: And Other Laws of Cyberspace, Version 2.0 (2006).

work developed in the previous section. It does so particularly examining the revised Product Liability Directive (PLD), and the AILD that was withdrawn by the European Commission.

Ultimately, choosing the right liability regime for AI value chains will depend on the characteristics of different value chains, the regulatory environment, and how risky the system at issue is. The main contribution of this Report is that it offers criteria to think about these questions. It also suggests that further empirical analysis on the way liability is distributed and enforced via private ordering, and on the costs and ease of proving fault in present liability may be important to establish whether regulatory interventions are needed, and to design them effectively.

How to read this report

This document is intended to accommodate a wider audience but is also specific and technical in the policy conversation it is engaging in. For this reason, the initial sections are spent developing certain important contexts on the AI value chain, and the core of the argument starts in Sections 3 and 4. Consequently,

- **If you are not a seasoned AI researcher or expert:** Sections 1-3 will walk you through the foundational concepts needed to grasp the main framework developed later in the Report.

- **If you are a researcher or policymaker working on AI**: You can skip Sections 1-3. These sections cover basic AI accountability challenges you're likely already familiar with. Focus on Sections 4 and 5, which provide the substantive analysis of AI liability along the value chain and core arguments.

- **If you are a lawyer,** you can skip Subsection 4.1, which covers basic liability concepts you are probably already familiar with. Dive back into Section 4.2, which starts developing the core argument and framework.

- **If you are mostly interested in the EU AI liability framework,** you may want to go straight to Section 5 which presents and analyses the EU AI liability framework. Reference other sections as needed, and specially Section 4, for more context and a deeper understanding.

- **If you are an AI policy enthusiast, don't skip the Annex**! Though not widely discussed, AI actors distribute and structure responsibilities and liability amongst them via contracts and licenses. The Annex provides a quick overview of how this happens.

Methodological clarifications

- This analysis is based on theoretical research about different liability systems. Regulators and other stakeholders should seek to complement such analysis with empirical analysis and real-world case studies of how specific liability systems handle AI to decide if reform is needed. They should compare the costs and benefits of changing and/or harmonizing liability laws to ensure AI accountability and victim compensation, versus other options like maintaining the current legal framework or introducing regulations.

- Most countries around the world have a rich tradition of tort or civil liability law. There are conceptual similarities in these regimes, but policymakers should consider national particularities when thinking about liability law reform for AI. For the purpose of this report, what follows mostly draws from the work of the Expert Group and the European Tort Law Group and their developed Principles of European Tort Law, which seeks to systematically understand liability law in European countries.[3]

3 PETL (n8)

01

Artificial intelligence accountability and liability today

01

Artificial intelligence accountability and liability today

Artificial intelligence is an umbrella term for related technologies and techniques. It encompasses techniques like machine learning and knowledge-based approaches, application areas like speech recognition, computer vision, and natural language processing, and the applications of these techniques in various domains.[4] The widely used OECD definition describes an AI system as:

> "...a machine-based system that, for explicit or implicit objectives, infers from the inputs it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptive deployment."[5]

AI systems are improving the efficiency and effectiveness of all sorts of products and services, and enabling new products and services in different fields, from personalized medicine to supply chain management. At the same time their deployment and adoption in certain sectors can pose important risks, specifically to safety and fundamental rights.[6]

When these risks materialize and harm occurs, a key question becomes who should be responsible. Under liability law, individuals are responsible for compensating victims for the harm their actions or omissions cause. Liability law is thus concerned with establishing when the burden of a loss must be shifted from the person who suffered it to the

4 OECD, "Explanatory memorandum on the updated OECD definition of an AI system," OECD *Artificial Intelligence Papers*, no. 8 (2024): https://doi.org/10.1787/623da898-en

5 *Ibid.*

6 *See* Marko Grobelnik, Karine Perset, Stuart Russell, "What is AI? Can you make a clear distinction between AI and non-AI systems?" *OECD AI Policy Observatory*, March 6, 2024

person who caused it.[7] To do so, liability law creates rules that establish when an action or inaction is legally considered to have damaged a legally protected interest. Usually, these rules require that a harmed individual shows in court that a legally protected interest (such as a fundamental right or a property right) has been harmed by someone else's actions or omissions and this has caused them damage.[8] However, even though AI systems are bound by these and other existing legal frameworks and principles — from data protection to consumer and antidiscrimination law — the application of these rules to situations that involve AI systems is challenged by AI systems' characteristics of opacity and complexity.

AI opacity refers to the difficulty that recipients of AI outputs may have in understanding how the system arrived at a particular output from the given inputs. For example, when a job application is automatically rejected by an AI system based on the resume submitted by an applicant, different circumstances may affect how easy it is for the applicant or even the human resources team involved to understand why they were rejected. The AI system can be opaque for different reasons: as a result of corporate secrecy, if the potential employer doesn't disclose the selection criteria and the procedures, technical or not, through which a decision is made (this form of opacity has less to do with the algorithm, and more to do with the transparency of the potential employer). AI systems can be opaque however, also as a function of technical literacy, if the people involved, like the applicant and the human resources team, have little understanding of how an algorithm works or may not be able to read basic code.[9]

Importantly, however, certain AI systems, like those that use machine learning, may also be opaque as a result of their technical complexity. Complexity refers to the fact that the outputs of AI systems arrive in non-linear ways, that is a form of logic that is not necessarily intelligible to humans. This occurs because these algorithms learn and process vast amounts of data, in ways that are impossible to replicate for the human mind. Even the builders of such AI systems might not be able to fully explain why a particular output was produced.[10] These types of opaque AI systems are often colloquially referred

7 *See e.g.* Ronald Coase, "The Problem of Social Cost", *Journal of Law and Economics*, 3, October 1960.

8 This varies in each domestic regime. The European Group on Tort Law explains that damage requires material or immaterial harm to a legally protected interest and their protection varies on its nature and value. Typically, life, bodily or mental integrity, human dignity, and liberty enjoy the highest protection. Property and property rights enjoy extensive protection, whereas the protection of pure interests in contractual relationships is more limited in scope. European Group on Tort Law, *Principles of European Tort Law* (Wien: Springer, 2005), Article 1:102(4) (*hereinafter* PETL).

9 Jenna Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms." Big Data & Society, 3(1), 2016.

10 European Commission, "Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) Explanatory memorandum", COM/2022/496 fina. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0496.

to as “Black Boxes”.[11] In these instances, even if the general logic of the algorithm is comprehensible, exactly how it reaches one decision rather than another may be impossible to understand.[12]

AI systems can also be complex in the sense that the tasks needed to complete the development and deployment of an AI system — from problem definition to data collection, labeling, cleaning, model training, fine-tuning, and testing and deployment — often involve many actors in a value chain.[13] This is often referred to as the “problem of many hands”.[14]

The fast adoption of AI across industry, and their opacity and complexity, have led regulators from around the world to adopt or consider regulations, soft law documents to improve their accountability because when harm occurs it is otherwise hard to know exactly what went wrong.[15] These regulations, for example, delineate requirements for deploying AI systems in especially sensitive contexts, or establish best practices and principles to test and develop AI systems to guarantee their safety. A notable example of such regulatory initiatives is the EU’s AI Act.[16]

Liability law is, however, an alternative or a compliment to regulation to create incentives for AI Actors to take measures of care and prevent harm.

When harm occurs, however, victims may still find it difficult to establish what led to the harm and what was the role of the different actors involved in its lifecycle, in its occurrence. Indeed, under most current national liability rules, and especially those relying on fault or negligence, victims of harm involving an AI-enabled product and service need to prove a wrongful action or omission by the person who allegedly caused the damage. AI’s characteristics, however, make it exceedingly difficult and expensive for victims to identify the liable party or parties and prove those requirements. Liability law alone may serve as a powerful tool to enhance AI safety. Indeed, if AI actors know they will be held respon-

11 IBM, *What is black box AI?* 29 October 2024 https://www.ibm.com/think/topics/black-box-ai

12 Burrel (n9)

13 Ian Brown, “Expert explainer: Allocating accountability in AI supply chains”, Ada Lovelace Institute, June 29, 2023 https://www.adalovelaceinstitute.org/resource/ai-supply-chains/

14 *See e.g.* Helen Nissenbaum, “Computing and Accountability”, *Communications of the Association for Computing Machinery*, 37(1), 1994; Batya. Friedman, “Moral Responsibility and Computer Technology”, *Institute of Education Sciences*.

15 *See* e.g. Algorithmic Accountability for the Public Sector – Report, *AI Now Institute, Ada Lovelace Institute and Open Government Partnership*, August 17, 2021. Available at: https://ainowinstitute.org/publication/algorithmic-accountability-for-the-public-sector-report

16 In the US, a notable example is the Artificial Intelligence Risk Management Framework (AI RMF 1.0) developed by the National institute of Standards and Technology https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

sible for harm, they will have strong incentives to prevent its occurrence.[17] Thus, regulators and legal researchers are considering whether, and if so how, liability rules should be applied or adapted to AI systems.[18]

This policy Report focuses on the challenge of assigning liability along the AI value chain, that is, across the many actors that often participate in the design, development and deployment of AI systems.

17 *See e.g.* Steven Shavell, *Liability for Harm versus Regulation of Safety* Working Paper No. 1218 National Bureau of Economic Research 1983. Available at: https://www.nber.org/system/files/working_papers/w1218/w1218.pdf

18 European Commission, Liability Rules for Artificial Intelligence (n.d) Available at: https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en; "ALI Launches Principles of the Law, Civil Liability for Artificial Intelligence", American Law Institute, October 22, 2024. Available at: https://www.ali.org/news/articles/ali-launches-principles-law-civil-liability-artificial-intelligence?_zs=3jEkb1&_zl=npEr9&utm_source=Informz&utm_medium=Email&utm_campaign=News&utm_term=ALI-Projects

02

The problem of many hands in AI accidents and different kinds of AI value chains

02

The problem of many hands in AI accidents and different kinds of AI value chains

The AI lifecycle consists in different stages of AI development and deployment: designing an AI system, collecting and processing data (which includes cleaning and labeling the data), building the actual model, verifying and fine tuning the model, deploying the system, and then monitoring.[19]

The multiplicity of actors involved in AI value chains represents a challenge for liability law because it is often difficult to know what each actor's responsibilities should be and how they contribute to AI outputs. In accidents that involve an AI system, it is thus difficult to identify the cause, and often the multiple mistakes or negligence, that led to the accident. This is a function of, and aggravated by, the opacity and technical complexity that characterizes AI systems.

AI value chains can have many forms, however. An AI system may be fully developed and deployed in-house in a firm or government agency. From a liability perspective, such a value chain does not present a problem for assigning liability. If the firm that develops the system does not take the right measures of care and it itself suffers the losses, it will have to bear them itself, as there is no external actor to assign liability to. If the victim is external, the liability will necessarily lie on the sole actor who developed and deployed the AI system.

19 OECD, *Recommendation of the Council on Artificial Intelligence,* OECD/LEGAL/0449; Advancing accountability in AI: governing and managing risks throughout the lifecycle for trustworthy AI", *OECD Digital Economy Papers*, No. 349, February 2023.

Many commercial AI systems, however, involve two or more actors:

Two-actor AI value chains may involve an AI system that is developed by a firm and is adopted and used by another firm. This is the case, for example, of a startup A that develops an AI model for predictive maintenance in manufacturing and sells or licenses that model to another firm B that uses the model in their operations. If an accident occurs in the operation of firm B, A and B will have to establish whether the accident was A or B's responsibility, and to what extent, to determine who has to pay for damages.

Often, however, AI value chains involve three or more actors: Microsoft Copilot, used by many businesses for a variety of tasks such as visualizing data or managing projects, integrates OpenAI's GPT model. Businesses are "users," Microsoft is the "provider" of the AI system, and OpenAI is the "developer." Other actors may also be involved: OpenAI may rely on third parties for collecting or labeling data,[20] and there is often an end user at the end of the value chain, such as the employees that use Copilot.

What follows is a simplified description of the key roles in an AI value chain. For clarity, this description adopts the definitions used by the European AI Act for two key roles in the AI supply chain: provider and deployer.

- **Suppliers of components:** Provides the necessary components and infrastructure required to build and train AI models. This can include hardware (like GPUs), datasets, and cloud computing resources.

- **Provider:** Creates the AI model or algorithm. Generally, creating a model involves research, design, coding, training the model with data, and continuous improvement. The AI Act defines an AI provider as the actor that develops an AI system and "places it on the market under its name or trademark."[21]

- **Deployer:** the actor that takes the AI model created by the Provider and uses it under its own authority and does so professionally.[22] They package, distribute, can

20 Billy Perrigo, "Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic", *TIME*, January 18, 2023, 7:00 AM EST. Available at: https://time.com/6247678/openai-chatgpt-kenya-workers/; LAION Roars: The story of LAION, the dataset behind Stable Diffusion", *The Batch,* June 7, 2023, Available at https://www.deeplearning.ai/the-batch/the-story-of-laion-the-dataset-behind-stable-diffusion/?ref=dl-staging-website.ghost.io.

21 European Union, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance) (*hereinafter* AI Act), Article 3(3).

22 AI Act (n21) Article 3(4)

fine-tune, and add support and additional features or integrate it into their own product or service. Note that an AI value chain may have multiple deployers in this sense: Microsoft is a Deployer of ChatGPT, a model developed by OpenAI, but a firm that uses Copilot may also be considered a deployer under this definition.[23]

- **End-user or operator:** Final user of the product licensed, or deployed by the deployer, such as the employee using Copilot.

Scholars have noted that AI value chains are complex because the actors and data flows of these value chains are in constant motion and that the responsibilities of different actors to assess and mitigate risks often overlap.[24] Take as an example of the type of liability question at issue the case of a chatbot in Air Canada's website that misled a client, telling them inaccurate information about the airline's bereavement travel policy: after Jake Moffatt's grandmother passed away, he needed to quickly book a flight from Vancouver to Toronto. He consulted Air Canada's chatbot about their bereavement fare policy before booking and the chatbot told him he could book a regular ticket and apply for a bereavement fare refund within 90 days. However, Air Canada's actual policy, to which the chatbot linked, specifically prohibited retroactive bereavement fare adjustments after booking.[25]

When Mr. Moffatt tried to get the refund as advised by the chatbot, Air Canada rejected his request. Despite showing them screenshots of the chatbot explicitly stating he could submit the ticket for a reduced bereavement rate within 90 days using their refund form, the airline refused to honor this. They claimed that Mr. Moffat should have known better because the chatbot had included a link to the correct policy elsewhere in the conversation.[26]

Once in court, Air Canada also alleged that it couldn't be held liable for information provided by one of its chatbots, and suggested that the chatbot was a separate legal entity, responsible for its own actions. The Judge reasonably disagreed. It found that Air Cana-

23 Article 25 of the AI Act establishes that when deployers put their name or trademark on a high-risk AI system, or make a substantial modification to it, they will be subject to the same risk-mitigation measures of AI providers of high-risk systems. *See* Section 5.

24 Brown (n13); Jennifer Cobbe; Michael Veale; Jatinder Singh, "Understanding Accountability in Algorithmic Value Chains", *ACM Conference on Fairness, Accountability, and Transparency*, 2023: https://arxiv.org/pdf/2304.14749

25 *See* Ashley Belanger, "Air Canada Has to Honor a Refund Policy Its Chatbot Made Up", *WIRED*, February 17, 2024, 12:12 PM. Available at: https://www.wired.com/story/air-canada-chatbot-refund-policy/; Canada, Civil Resolution Tribunal, "Moffatt v. Air Canada", SC-2023-005609, Small Claims, February 14, 2024. Available at: https://www.canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html

26 *See* Belanger (n22); *Moffatt v. Air Canada* (n22)

da "did not take reasonable care to ensure its chatbot was accurate."[27] It also didn't find there was a significant difference between Air Canada's responsibility for the information provided by the chatbot, a component of its website, vis a vis any other component of the website.[28] Lastly, the judge found that Mr. Moffat had reasonably relied upon the chatbot providing reliable information, and that there was no reason for him to double check the information provided by Air Canada's website.[29] Air Canada was ordered to honor the chatbot's made up bereavement policy, and reimburse part of the tickets to Mr. Moffat.[30]

The judge's decision seems reasonable, as corporations are generally held liable for the dependants and by the information they tell their clients, so long as they take reasonable care to ensure their representations are accurate and not misleading. Here, the judge considered that there was nothing apparently different even if a chatbot was involved.[31] But imagine the case is slightly different. Imagine the chatbot is provided by a third party company — as Air Canada may have implied — and it is somehow branded as such on the website. This is, thus, at least a two-actor AI value chain. Let's take into consideration, in addition, the fact that chatbots powered by large language models can indeed make information up — these are called "hallucinations".[32] To deploy a chatbot safely (or as accurately as possible), AI developers and deployers should program guardrails into the system to ensure it follows policies and ethical guidelines, and must fine-tune it to adapt to new data or interactions, but this rarely ensures full accuracy.[33] In such a scenario, many questions arise: Is it Air Canada's obligation or the chatbot developer to ensure the closest to accuracy? Or is it both? If such a chatbot (developed and branded by a third party) misrepresents information to a client, should Air Canada be held liable, the AI developer, or should the AI developer compensate Air Canada for having been held liable? Or is misrepresentation by a chatbot something that, in the long run at least, end-users should assume because the average consumer should know better?[34]

27 *Moffatt v. Air Canada* (n25) at 27.

28 *Moffatt v. Air Canada* (n25) at 27.

29 *Moffatt v. Air Canada* (n25) at 28.

30 *Id.*

31 *Id.*

32 Michelle Hawley, "Exploring Air Canada's AI Chatbot Dilemma", *CMSWIRE*, April 2, 2024. Available at: https://www.cmswire.com/customer-experience/exploring-air-canadas-ai-chatbot-dilemma/

33 *Id.*

34 Some scholars have argued reasonably argued against this last statement, *see* Sandra Wachter, Brent Mittelstadt, and Chris Russell, "Do large language models have a legal duty to tell the truth?", Royal Society Open Science, January 31, 2024, Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4771884, In the professional context, however, lawyers using chatbots that misrepresent information (or make it up) have been told by courts that they should know better. See Sara Merken, "New York lawyers sanctioned for using fake ChatGPT cases in legal Report", *Reuters,* June 26, 2023, 10:28 AM GMT+2. Available at: https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases-legal-Report-2023-06-22/ .

This is the type of liability question that arises when multiple actors are part of an AI value chain. There is a question about what are the reasonable expectations society should have from each actor, including end-users. There is also a question about the mitigation measures actors should take, and whether everyone has sufficient incentives to take them. Some actors may not have such incentives, for example, if they evaluate as low the likelihood that they may have to pay if harm occurs (that is, there is a lack of adequate enforcement), or if they have inadequate information about the likelihood and gravity of harm that their actions and inactions may cause. This may be the case, for example, of data providers that do not invest enough in data cleaning or privacy measures, or the developers of models that do not test enough for bias. The lack of incentives to take care, and the resulting lack of care, however, ultimately increases the probability of harm occurring.

03

Roles and responsibilities along AI value chains and how they are structured

03

Roles and responsibilities along AI value chains and how they are structured

Actors in an AI value chain often create responsibilities and distribute liability amongst them through "private ordering", that is, via contracts amongst them, or their licenses, and terms of service. The architecture of a particular system, and the market relationships amongst them may also play a role in helping define who is in charge of what and who will be held liable in case of harm.[35] These factors are important because they shape what is and can be expected from the different actors in the AI value chain and consequently help determine who may be liable for AI-related harm, as the law typically mandates that individuals who breach their responsibilities are held accountable and must pay for any resulting damage.

What follows is a simplified explanation of how these different mechanisms structure the roles and responsibilities along the AI value chain.

3.1 How technical factors and release decisions structure roles and responsibility along the AI value chain

The design and interface choices made by AI developers shape what other actors downstream can or cannot easily do with an AI system. These technical factors and architecture features structure AI value chains that are possible for different actors in an AI value chain. OpenAI, for example, allows for the fine-tuning of its GPT models, but it does so through an application programming interface (API) that makes specific technical decisions that limit how downstream developers and companies can use it.

35 This framework, which holds that responsibility and liability are structured and governed by architecture, the market, law and norms, echoes Larry Lessig's pathetic dot theory. Lessig (n3).

These include, for example, limiting the amount of text developers can enter into GPT, or OpenAI's design choices about the format in which responses come out. Thus, using GPT, a developer can build chatbots, virtual assistants, or other AI applications, but must work within the constraints of the OpenAI API. For instance, the API limits the maximum number of tokens (linguistic units, like words or parts of words) that can be processed in a single request, and it requires input (prompts) and output (model-generated response) to be in specific format.[36] These architectural decisions by OpenAI shape how developers can implement and integrate GPT into their projects.[37]

On the other hand, "open-weight" models like Llama disclose the model parameters and allow users to freely tinker with the model and fine-tune it for specific tasks. This, in some cases, gives developers more freedom in developing their own customized GPTs compared to the model preferred by OpenAI.[38]

The relevance of these architectural choices is significant because they enable and constrain different kinds of uses. Recall that liability law mostly relies on establishing whether a given actor acted as expected from them when harm occurred. Think of an analog version: home appliances, like blenders, are designed for home uses, such as making soups and thus blending relatively soft materials (vegetables, fruits, etc.). It is very difficult for a consumer to, for example, cook something using just a blender, because the tool is just not designed for it. If an accident occurs when a consumer is trying to cook something with a blender, or trying to blend something that is not food — a type of stone, for example — then it is most likely the consumer who is not behaving as reasonably expected. If a harm occurs when the consumer is using the blender just as instructed, however, we may suspect the blender manufacturer.[39] Design and architecture frame what is possible, easier, and reasonable to do with a given tool or model, architectural choices made by an AI developer frame what is reasonable to expect from different actors in the value chain.

36 The format is JSON formatted payload

37 Note that when an API user "breaks" into the system and does something that OpenAI had tried to prevent, and causes harm, that harm is, most likely, the user's responsibility.

38 Parth Nobel, Alan Z. Rozenshtein, and Chinmayi Sharma, "Open-Access AI: Lessons From Open-Source Software", *Lawfare*, October 25, 2024, 10:00 AM. Available at: https://www.lawfaremedia.org/article/open-access-ai--lessons-from-open-source-software

39 Take as an AI example a model is designed by the AI developer (A) to analyze medical images for diagnostic purposes of a particular disease and is licensed to another entity (B) via an API. Imagine a first scenario, in which B installs the API and uses it as instructed, but something goes wrong and harm occurs because the model misidentified the condition, and some patients went into a surgery they didn't need. In this case, suspicion may fall upon A, given that A designed the model and the harm occurred while the model was used as expected.
But let's consider a second example. Imagine that B decides that it wants to use this same model to diagnose a somewhat similar, but albeit different condition. When doing so the model misidentifies the new condition, and some patients go into a surgery they didn't need. Now, the suspicion falls on B because B used the API for purposes other than its intended design, and ignoring usage guidelines.

Design and architectural choices are not the only way, in which developers can impose constraints on what downstream users can do with their models. Even developers releasing open models may impose certain constraints through licensing terms (see below).

3.2 How contracts, licenses and terms of service govern roles and responsibility along the AI value chain

Actors in an AI value chain regulate the obligations and responsibilities of the different actors of an AI supply chain through their terms and conditions, licenses, and contracts. This is called "private ordering". These contracts, for example, allocate responsibility and liability amongst them the different actors for specific damages — such as for outputs that infringe copyright [40].

Scholars have shown that the terms of service of general-purpose models — like Open AI's GPT models, Meta's Llama, or Anthropic's Claude — tend to defer liability for outputs down the value chain.[41] For example, OpenAI's terms of use specify that their services are provided "as is" without any warranties, and users accept all risks associated with using the outputs. Additionally, users or providers must indemnify OpenAI against costs and liabilities arising from third-party claims.[42] Similarly, Llama's Community License Agreement grants users a non-exclusive license to use and modify the Llama Materials but includes a disclaimer of warranty, indicating that users use Llama at their own risk.[43]

In the B2B sector, however, companies tend to allocate liability more evenly, establishing specific obligations, and establishing who should take responsibility for specific damages when they occur. These are important institutions, because even if victims of harm can bring a claim against all the actors that potentially caused it regardless of whether they have a contractual relationship with them, courts will look at these arrangements to understand the expected role of the different actors.[44] They will generally also uphold the contractual provisions agreed upon by the different parties of the AI value chain unless these are considered to go against the law and public order.

40 Lilian Edwards, "Private Ordering and Generative AI: What Can We Learn from Model Terms and Conditions?", *forthcoming in the Zou et al. eds Cambridge Research Handbook on Generative AI and the Law (CUP 2025)*, November 19, 2024, at 13. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4771884.

41 *Id.*

42 OpenAI Terms of Use, December 11, 2024. Available at: https://openai.com/policies/row-terms-of-use/

43 "LLAMA 3.3 Community License Agreement", *Llama*, December 6, 2024. Available at: https://www.llama.com/llama3_3/license/

44 This wasn't always the case. In the United States, for example, the famous case that first allowed this was *MacPherson v. Buick Motor Co*, in 1916. Previously the general rule had been that only the parties of a contractual relationship with a product's manufacturer could sue for the products malfunctioning and the damages they caused. *See* Marler Clark, An Introduction to Product Liability Law. Available at: https://marlerclark.com/pdfs/intro-product-liability-law.pdf

Annex I further explores the ways in which different AI actors govern AI liability, absent regulation, through their terms and conditions, licenses, and contracts regulate the obligations and responsibilities of the different actors of an AI supply chain and how they would interact with existing liability law in a given litigation.

3.3 How market power influences roles and responsibility along the AI value chain

Lastly, economic factors govern and structure AI value chains, by allowing large developers and providers to leverage economies of scale to develop and provide access to models that would be very hard to develop for small companies. Economic factors also lead to the increasing consolidation of the AI value chain around certain systemically important AI developers and service providers, such as Microsoft, OpenAI, Amazon Web Services, and Meta. Importantly, this economic consolidation renders these actors as important nodes in value chains to check for the suitability and safety of certain components, which they most likely do. Regulators could consider the merits and costs of using economic consolidation around these key actors when trying to find points of interventions and responsibility on the AI value chain.[45]

45 *See* Jennifer Cobbe, "Governance and Interdependence in Data-Driven Supply Chains", *forthcoming in Fleur Johns, Gavin Sullivan & Dimitri Van Den Meerssche (eds), Global Governance by Data: Infrastructures of Algorithmic Rule*, April 2, 2024.

04

Liability rules as applied to the AI value chain

04

Liability rules as applied to the AI value chain

Liability law is concerned with establishing the rules that determine when the burden of loss must be shifted from the person who suffered it to the person who caused it, and when an action or inaction is legally considered to have damaged a legally protected interest.[46]

The previous section highlighted the crucial roles that technical factors, economic factors, and legal documents play in assigning responsibility and liability within an AI value chain. Courts may consider design features of market power when establishing what is reasonable to expect from whom, yet they will uphold legal agreements, such as those presented in the previous section and Annex I, if they comply with established liability principles and existing regulations. Liability law is the institution that ultimately determines who is held responsible when harm occurs.

If liability reform is deemed necessary, this section offers an overview of the main policy choices at stake when designing liability regimes for multiparty situations, such as the AI value chain: whether fault-based liability or strict liability should be chosen, whether parties should be jointly and severally liable or only severally liable, and who should bear the burden of proving that something went wrong.

4.1 The form of liability: Fault liability or strict liability?

Liability law determines who is legally responsible for causing harm. There are two main standards for establishing this responsibility: fault-based liability and strict liability.

46 Coase (n7)

- **Fault-based liability** means an actor is considered liable when their negligent action or omission—that is, failing to meet the expected standard of care—leads to harm. They are held accountable because they didn't act as carefully as they should have.

- **Strict liability** holds actors responsible for harm caused by their activities or products regardless of negligence. Even if they took all necessary precautions and acted with due care, they are still liable if their actions result in harm.

In most countries, fault-based liability is the general rule. Strict liability is typically reserved for activities where significant risks remain even when all precautions are taken. These are generally referred to in the liability literature as "dangerous activities." Examples include handling hazardous chemicals or operating certain types of heavy machinery, where the inherent risks can't be completely eliminated.[47]

a) Fault-based liability

Fault-based liability is the general rule in many places around the world. The main reason is that it is considered a fairer standard, as it requires proving that the defendant was negligent, and thus only actors that act carelessly are held accountable for harm. Consequently, in cases where no one really made a mistake, victims of harm bear the loss.

From an incentives perspective, fault-based liability encourages actors to be careful in their activities (developing AI models, deploying them, and so on), as if they are able to prove that they took due precautions they will not be held liable even if an accident occurs. It also encourages potential victims of harm to also be careful, as they have an interest in preventing accidents.

However, there are different challenges associated with fault-based liability: An important challenge concerning AI systems and the AI value chain is proving that someone met, or didn't meet, the expected standard of care.[48] Courts and claimants may find it hard to identify if enough precautions were taken, may lack information to correctly assess the level of care taken or may not have full details of certain measures of care.[49] Similarly, proving causality — that an actor's "faulty behavior" caused a harmful outcome — may be

47 *See* Steven Shavell, The Mistaken Restriction of Strict Liability to Uncommon Activities, Journal of Legal Analysis, Volume 10, 2018 [*hereinafter* The Mistaken Restriction of Strict Liability].

48 Buiten et al. (n2)

49 Shavell, The Mistaken Restriction of Strict Liability (n48) at 13.

difficult due to AI systems' complexity.[50] Thus, litigation in fault-based systems may be long, expensive, and uncertain. A strict-liability regime may offer more legal certainty.

b) Strict liability

Unlike fault-based liability, strict liability means that victims only need to prove that the defendant's actions caused the harm, regardless of whether the defendant took optimal care.

This is generally an exceptional regime, reserved for activities that are considered to remain dangerous even when precautionary measures are taken. Yet, the activities covered by strict liability vary widely across different countries.[51] In many continental law systems, for example, individuals are held responsible for the actions of the "things" or people under their control, like employees or vehicles they operate. This is known as vicarious liability.[52]

Strict liability has several advantages, especially when applied to dangerous activities.[53] So-called dangerous activities are activities that remain risky, even when appropriate care is taken. In such cases, strict liability encourages actors to take appropriate safety measures because they have a strong incentive to prevent harm—they will be held liable regardless of their level of care. Additionally, there are important advantages for victims under a strict liability regime, because it exempts them from proving that the defendant did not take adequate measures of care. Victims of harm would only have to prove that the defendant's actions led to harm.[54]

Importantly, it motivates actors to moderate the levels of activities — that is, to do less of them, only as much as makes economic sense. Since they bear the potential costs of liability even when careful, they will only engage in activities that are economically viable for them.[55] An additional advantage of strict liability is that it is much easier for victims of harm to prove their case, because they only need to show that the harm was caused by the defendant's actions, not that the defendant was negligent.

50 Buiten et al. (n2)

51 Strict Liability", *Legal Dictionary*, March 13, 2016. Available at: https://legaldictionary.net/strict-liability/.

52 PETL (n8) Article 6:102.

53 Shavell, The Mistaken Restriction of Strict Liability (n48)

54 It is important to consider that strict liabilities are frequently paired with liability caps or other limitations to balance the heightened risk for those benefiting from the technology. *See* Directorate-General for Justice and Consumers (European Commission), Liability for artificial intelligence and other emerging digital technologies (2019), at 23. Available at: https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en

55 Shavell, The Mistaken Restriction of Strict Liability (n48)

There are several limits and downsides with choosing strict liability as a liability regime. First, strict liability may only be socially useful when activities are dangerous even after precautions are taken. When the activities at issue aren't dangerous, proper levels of care are achieved under the fault regime, which may be less costly for potential injuries.[56]

Second, strict liability does not encourage victims or others under the defendant's control to take precautions. They might rely on the fact that the defendant will be liable regardless.[57] This issue can be partially addressed by allowing the "concurrent fault of the victim" defense, where victims are responsible when their negligent behavior contributes to harm.[58]

Third, because strict liability increases the private costs of the activity and has thus an effect on the levels of activity, strict liability can in theory discourage beneficial activities because it imposes additional costs without accounting for their benefits. For example, AI applications like autonomous vehicles and diagnostic tools can potentially offer significant societal benefits, and not using them could lead to missed opportunities. For similar reasons, strict liability may deter entrepreneurs from pursuing innovative AI projects or securing funding.[59] On the other hand, some authors have pointed out that strict liability offers higher legal certainty, as what exactly constitutes fault may be hard to establish.[60]

Lastly, it is worth noting that strict liability may lead to a higher demand for AI liability insurance from both cautious actors and potential victims seeking coverage for accidents. While developing such an insurance market could have positive effects by providing better risk coverage, it could also create new burdens, like the administrative and social costs of insurance markets and insurance disputes. [61] Scholars and regulators

56 Shavell, The Mistaken Restriction of Strict Liability (n48)

57 Buiten et al. (n2)

58 Note however, that there may always be a tradeoff regarding whose optimal behavior should be encouraged: Under strict liability with a contributory negligence defense, victims are motivated to avoid behaviors that would clearly be deemed contributory negligence, such as running into heavy traffic. However, they might not take other precautions, like wearing reflective clothing at night, if these are not typically considered in contributory negligence cases. In contrast, under the negligence rule, victims are encouraged to engage in risky activities only to a socially appropriate extent. This mirror effect shows that the risk-reducing benefits of the negligence rule for victim behavior might be more significant than the advantages of strict liability for injurer behavior. It may be that often, however, it is more important to control the behavior of the injurer than of victims. *See* Shavell, The Mistaken Restriction of Strict Liability (n48)

59 *See* Buiten et al. (n2)

60 *See* the Directorate-General for Justice and Consumers (n55) at 29 arguing that "this allegedly chilling effect of tort law is even stronger as long as the question of liability is entirely unresolved and therefore unpredictable, whereas the introduction of a specific statutory solution at least more or less clearly delimits the risks and contributes to making them insurable."

61 Steven Shavell, "Liability for Accidents", *School of Law, Harvard University, and National Bureau of Economic Research* [*hereinafter* Shavell, Liability for Accidents] Available at: http://www.law.harvard.edu/faculty/shavell/pdf/07-Shavell-Liability%20for%20Accidents-Hdbk%20of%20LE.pdf

considering different liability regimes may want to consider the costs and benefits of such a development.

c) Product liability

Product liability is a specific form of liability that covers producers of material products used by consumers, such as smartphones or microwaves, though recently, the EU reformed its product liability legal framework to also include software.[62] Under product liability law, a producer is responsible for damages to life, health, and property caused by a defect in their product.[63]

Though product liability is in principle a form of strict liability, many scholars recognize that it has elements of fault-based liability, depending on how "defect" is interpreted. A defect is generally understood as a deviation from the safety which a person is entitled to expect from that product or service and the state of scientific knowledge at a given time.[64] As Buiten explains, "[i]f courts assume a defect easily, product liability tends towards strict liability. Conversely, if proving a product is defective is difficult, product liability resembles a negligence rule." [65]

In practice, this requires proving whether proper measures of care were taken to prevent a defect. Developers can also defend themselves by proving that the state of the art in science and technology could not have detected the defect when the product was put in circulation.[66]

Besides the benefits and difficulties of using fault-based liability already discussed, product liability law is generally considered a good alternative for instances where there is one actor in the value chain that is in a better position to prevent harm. This is typically because they have better information or control over the situation, and they are

62 *See Consolidated text: Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.*

63 *Id.*

64 Andrea Bertolini, "Artificial Intelligence and Civil Liability", *Study requested by the JURI committee, Policy Department for Citizens' Rights and Constitutional Affairs*, July 14, 2020. ; PLD Article 11(e).

65 Miriam C. Buiten, "Product liability for defective AI", European Journal of Law and Economics, February 27, 2024, Volume 57, pages 239–273 https://link.springer.com/article/10.1007/s10657-024-09794-z.

66 *See* Philipp Hacker, "Proposal for a Directive on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence: Complementary Impact Assessment", *EPRS | European Parliamentary Research Service* (*hereinafter* Hacker, *Impact Assessment*)

able to influence the manufacture and design of the product.[67] For example, the provider of a particular AI system that has significant market power or technical capacity may be in the position to influence actors from the value chain — such as component manufacturers, developers and data labelers. Following this rationale, the actor best placed to guarantee — and be responsible for — the safety of a product is also known as "the cheapest cost avoider." [68]

There may be instances, however, where the cheapest cost avoider is not the provider of the system. Take, for example, the case where harm is caused by a defect in a component that is very hard to detect. If the AI provider is held liable, and if the product manufacturer knows that the provider is most likely to be held liable, and that the defect is hard to detect, then the product manufacturer has less incentives to take all the due measures of care. [69]

In some jurisdictions, courts may be empowered to identify cheapest cost avoiders. Since identifying the cheapest cost avoider from the outside is hard, some scholars have also proposed that the law assign liability to a specific type of party — such as AI providers — and allow that party to negotiate with others (such as deployers) to shift the liability to the actual least-cost provider via contractual ordering.[70]

4.2 Dividing liability or not amongst multiple potential injurers

A central challenge in designing liability frameworks for the AI value chain is figuring out how to distribute liability among the many involved actors and create incentives that encourage everyone to take proper care. When multiple actors need to take precautions, the second policy question is how to divide liability amongst them: will all the responsible actors be liable for all of the harm (that is, can the victim sue any of them and recover the entire loss)? Or should each actor only be held liable for the extent to which they are responsible? [71]

67 This can also be described as a broadened interpretation of the principle of vicarious liability. *See also* Catherine M. Sharkey, "A Products Liability Framework for AI", *Columbia Science and Technology Law Review*, Vol. 25, No. 2, 2024, March 26, 2024. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4773874.

68 Sharkey (n68)

69 William M. Landes and Richard A. Posner, "Joint and Multiple potential injurers: An Economic Analysis," The Journal of Legal Studies 9, no. 23 (1980): 535. Available at: https://www.jstor.org/stable/724003.

70 Buiten et al. (n2)

71 This is unless we are under a product liability regime, where the producer is the relevant actor to sue.

Legal systems vary in how they distribute liability amongst multiple potential injurers. In European countries, when two or more actors cause harm together, the general rule is that they are held under so-called solidarity liability (also known as joint and several).[72] Joint and several liability means that victims can demand full compensation from any of the actors that contributed to harm, regardless of the degree to which they caused it. The paying party can then seek reimbursement from other responsible actors based on their share of responsibility.[73]

So-called several or divisible liability is applied when it is reasonable to assign only part of the damage to each actor involved.[74] In such a case, each injurer is only liable for the part of the damage attributable to them.[75] This is the case, for example, where one activity led to damage, and a subsequent activity aggravated the damage: imagine an automated car swerves on a curve and crashes into an old, dry tree, causing the tree to fall and damage a fence. The initial damage is caused by the car, while the subsequent damage to the fence is caused by the tree. In this case, the car's owner is responsible for the initial damage, and the tree's owner is liable only for the damage to the fence. [76] It is, however, generally up to the injurer to prove that the damage is severable.[77]

TABLE 1	
Joint and several liability	**Several liability**
The victim can seek compensation from every injurer for the totality of the damage.	The victim can seek compensation from each injurer for the part of the damage they contributed to.

Table 1: Joint and several liability vs. several liability

As in section 4.1, the choice for several or joint and several liabilities creates different incentives: Several liability may be considered fairer, as parties will only be held liable for the harm they cause. However, establishing the extent to which each party should pay may be impossible or exceedingly difficult. It may also be difficult for victims to know

72 PETL (n8) Article 3:103(3).

73 If it is not possible to determine the relative responsibility of the persons liable, they are then to be treated as equally responsible. However, if one of the persons liable is an auxiliary of another one, the superior should be treated as sharing the whole responsibility. PETL (n8) Article 9.

74 It is, however, generally up for the potential injurer to prove that the damage is severable. In such a case, each potential injurer will only be liable for the part of the damage attributable to them. PETL (n8) Article 9:101.

75 *Id.*

76 PETL (n8) Article 3:104 Depending on the system, the car owner — unlike the tree owner — may be held solely liable for the totality of damages.

77 PETL (n8) Article 9:101

whom to sue. Joint and several liability, on the other hand, allows victims of harm to sue any of the parties of the value chain who they think may be liable.

Scholars have also shown that, when combined with the fault standard, joint and several liability may create optimal incentives for parties to take care in situations where there are multiple potential liable parties. This occurs because joint and several liability increases the potential cost of liability for each of the actors, as they can be held liable for the whole potential damage, not for the part they may have contributed to. At the same time, any innocent actor will be able to avoid liability by showing that they did care.[78]

On the contrary, when combined with a strict liability regime, joint and several liability makes joint injurers distrust each other, leading to a 'race to the bottom' situation. This occurs because actors in a value chain will calculate as the cost of their activities the cost of care plus the cost of potentially being held liable.[79] Consequently, if one actor in the AI value chain has reasons to think that another actor isn't taking enough measures to prevent harm, they might be incentivized to also lower their own care measures as a way to save net costs. This happens because if the first actor doesn't take care, the risk of harm goes up, and the second actor might be held liable anyway. To avoid high costs, they might reduce their care efforts, leading to a free rider problem.

The following table summarizes the policy choices at issue when assigning liability along an AI value chain, and highlights some of the key elements of each choice where multiple parties may be responsible.[80]

78 Posner & Landes (n70)

79 In economic analysis of liability law, this is calculated as the potential cost of harm multiplied by the probability of harm.

80 Note that this table always assumes, as in continental countries in general, that the "concurrent fault of the victim" is an available defense.

TABLE 2		
Liability regime / Responsibility amongst multiple potential injurers	**Fault-based**	**Strict liability**
Joint and several	It is not very hard to establish what are the appropriate levels of care. Parties do not need to be able to trust that other actors in the value chain will take optimal care. They will take due care because the risk of carrying the totality of liability is important. However, they will continue to carry on with their activities while exercising care because they may be exempted from liability by showing that they did so.[81]	The activity at issue is still dangerous, even if precautions are taken. This may, however, significantly deter the levels of activity. Establishing to what extent each party is liable is exceedingly costly. Parties may be led to take suboptimal levels of care if they believe another party is not taking care, as if a damage occurs, they may equally be held liable (that is, there is a free-rider problem).[82]
Several	The activities at issue are net socially desirable. It is easy to establish to what extent at least one actor contributed to harm, for example, because they aggravated it.	The activities at issue are still dangerous, even if precautions are taken. It is not hard to establish what are the appropriate levels of care. It is easy to establish to what extent each contributed to harm, for example, because they aggravated it.
Cheapest cost avoider (as in the product liability regime)	There is one actor well positioned to mitigate harm because they have significant control over the value chain and the outcome of the AI system and quality information and/or can bargain with other actors in the value chain to apportion liability effectively.	

Table 2: Different configurations of liability regimes

81 In jurisdictions where parties may always recover damages from other potential injurers, this rule may still be suboptimal. *See* Buiten et al. (n2)

82 *Id.*

4.3 Information asymmetries and the burden of proof

The last policy question to consider relates to the burden of proof. When a harm occurs victims of harm need to prove the elements of liability: Under fault-based liability, one must prove that harm occurred and that a faulty activity caused it. As discussed already, this is hard to do because AI systems are complex and opaque and because determining what exactly should be the desirable safety measures may require a lot of expertise. Strict liability only requires proving that a harm occurred and that this was caused by someone's actions. This already simplifies the process for victims because they do not have to prove that the action was negligent.

Policymakers (and in some jurisdictions, courts) can, however, alleviate the burden of proof in two main ways:

1. **Shifting the burden of proof:** The liability regime (or a court) can require the potential injurer, rather than the victim, to prove a particular element. For example, a regulation could establish that when victims prove that an AI actor was negligent, it is assumed that this negligence led to harm unless the AI actor proves otherwise.

2. **Providing access to information:** When harm results from the actions or omissions of several actors, identifying the liable party can be difficult and costly. It is also difficult and constantly to prove that something went wrong. To address this information asymmetry, regulations or courts may require AI actors to disclose relevant information about the AI value chain or system documentation. The downside of these measures, however, may be that they incentivize frivolous claims. This might also compromise IP and other rights, and it may significantly shift the distribution of risks towards AI developers or manufacturers.[83]

4.4. Preliminary framework

Designing liability regimes for AI actors along the value chain requires carefully considering how different rules create different incentives. Important factors to consider are how easy it is to determine that appropriate care was taken, the moral hazard effects strict liability may have on victims and AI operators, the effects on innovation, the procedural costs of each rule, and whether there is one specific actor in the value chain that is the cheapest-cost avoider (and/or who can effectively bargain with the other actors).

83 Buiten et al. (n2)

The analysis above suggests that strict liability is the most efficient liability rule where only one actor is involved, where the concurrent fault of the victim defense is available, and when the main goal is to diminish the probability of harm. This would alleviate the burden of proof of the victim, advance legal certainty, and create incentives for everyone — AI deployer and potential victim, to take care.[84] To create incentives for victims to take care when that is a factor, the concurrent fault of the victim defense should always be adopted.

Choosing a strict liability regime as a baseline regime for AI systems, however, faces two main challenges. First, a strict liability regime may have undesirable effects on the levels of activity and incentives to participate in the market of certain AI actors. Second, a strict liability regime, when factoring in the ways in which AI actors in the value chain may strategize may contribute to actually diminishing the level of care of certain actors in the value chain. These two conclusions are explained as follows:

1. **Most AI systems are not inherently dangerous and all AI actors don't have sufficient control over the AI system.** Consequently, a strict liability regime could lead to over-deterring societally desirable AI developers and providers from participating in the market.

   Thus, a fault-based regime is in general more desirable as a baseline, unless the AI system is one where inherent important risks remain after care is taken, and having incentives to diminish the levels of the activity at issue is socially desirable.

   Where a particular jurisdiction already assigns a strict-liability standard to a particular activity (i.e. driving), however, it may make sense to maintain that standard for those activities.[85]

2. **Strict liability and joint and several liability may lead to a free rider problem:** If joint and several potential injurers under strict liability cannot trust that the other party is acting with sufficient care, they may have incentives to themselves not act with the appropriate level of care since they might be held liable regardless of what they do. This increases the likelihood of harm.[86]

   The free-rider problem may not exist, however, if the parties can trust each other's behavior, for example because they are contractually bound and have an interest

84 Posner & Landes (n70)

85 *See* Directorate-General for Justice and Consumers (n55)

86 In jurisdictions where parties may always recover damages from other potential injurers, this rule may still be suboptimal. *See* Buiten et al. (n2); Posner & Landes (n70)

in their ongoing relationship. In such an instance, a strict liability regime where all parties are held joint and several liable would also incentivize all actors to take care and will lead parties to internally penalize those who do not take appropriate precautions. The risk of over deterring certain actors, however, remains.

3. **Fault liability and joint and several liability create better incentives to take care for everyone in the value chain:** Where the contractual relationships are weaker, the free-rider problem may be addressed by the fault standard, as parties can be exempted from liability by showing due care, even if harm occurs. The challenge here, however, is the difficulty in establishing fault.

4. **When should several liability be preferred?** Joint and several liability seems to be desirable in instances where the degree to which different actions contributed to harm is hard to establish. Several liability is more desirable where this is easier to establish. However, it should be up to the potential injurer to prove that the damage is severable. [87]

5. **Regulation and liability regimes interact in complex ways.** It is important to consider how liability regimes will interact with regulation when regulation is in place. Liability law compliments regulation by providing an avenue for victims of harm to seek compensation when regulation is in place. Where regulation is in place, liability law ideally acts as a complement to regulation by, for example, encouraging AI actors to compensate for regulatory vacuums when "reasonable safety measures" are different than those required by regulation. Under a strict liability regime, liability law may create incentives for AI actors to exceed regulatory requirements (provided that they are not under a situation that incentivizes them to "free ride" on other AI actors).[88]

   This complementarity of regulation and liability is advantageous, because there is still important uncertainty about the nature and scale of many risks associated with AI systems, as well as what exactly reasonable care entails. At the same time, policymakers and researchers should carefully evaluate the costs and benefits of imposing a regulatory burden and a strict liability regime on AI providers subject to regulation. It may be that, in certain circumstances, AI Actors complied with all the regulatory requirements were, in fact, taking care and imposing on them a strict liability regime may unduly burden them and disincentivize their choice to

87 *See* PETL (n8) Article 9:101

88 *See* Steven Shavell, "On the Social Function and the Regulation of Liability Insurance," *The Geneva Papers on Risk and Insurance* 25, no. 2 (2000): 166-179, https://www.jstor.org/stable/41952522

participate in the market and disincentivize potentially beneficial applications. Consequently, it may be undesirable to extend strict liability to all the already regulated high-risk systems under the AI Act. Future research should determine if AI providers' activities remain risky when they follow regulations, and if using strict liability on regulated AI systems reduces their activity levels more than socially desirable.[89] In other circumstances, however, the inherent risks created by certain AI systems may be such that even when AI actors comply with regulation and they take due care they should be held liable if harm occurs, and that imposing strict liability does not create free-riding incentives for other AI actors.

Consequently, the conclusion here is that joint and several fault-based liability is the best baseline regime for assigning liability in the AI value chain. Indeed, where there are concurrent causes to an accident or causes that may be indistinguishable, fault-based liability combined with joint and several liability creates good incentives for all the actors in the value chain to take measures that will minimize the cost and likelihood of an accident, and it also facilitates the victim's compensation.[90] This baseline regime is best complemented with measures that alleviate the burden of proof of victims, such as shifting the burden of proof when AI systems are complex or enabling access to evidence.

As above, however, this baseline regime comes with important challenges. Establishing what are the optimal dimensions of care may be exceedingly difficult for victims and courts and the litigation costs (for private actors, but also society at large) may be significant.[91] In general, policymakers and scholars would greatly benefit from further empirical research that examines the costs and the ease or difficulty with which victims pursue liability claims and how these are shared or not along value chains.

Because these costs and challenges of such an arrangement may be high, however, it may be desirable that on a case-by-case basis, courts and regulators extend strict liability or product liability to situations where there are actors in the AI value chain that have important control over the AI outcome. This may be the case, for example, of AI systems meant for consumer use, or, where an AI value chain is such that there is one actor that can be considered to have significant control over the other actors and the risk.[92] In such

89 Philip Hacker, for example, makes a strong argument that strict liability should be adopted for high-risk systems as defined under the AI Act. Hacker, *Impact Assessment* (n67)

90 A similar conclusion has been reached by Buiten et al. (n2)

91 Shavell, The Mistaken Restriction of Strict Liability (n48)

92 *See* the Directorate-General for Justice and Consumers (n55) at 41 similarly concluding that strict liability is appropriate for technologies operating in "non-private environments and may typically cause significant harm." In such cases it should be assigned to the individual controlling and benefiting from the operation of emerging digital technologies. If there are multiple operators, such as a frontend operator who primarily decides on and benefits from the technology, and a backend operator who defines its features and provides support, liability should rest with the one exerting more control over the operational risks.

instances, and always maintaining the defense of the fault of the victim or end-user, subjecting this actor to a strict liability regime (or a quasi-strict liability regime, like product liability), and allowing them to divide liability amongst providers via private ordering may end up being the more efficient rule.[93] Additionally, some of this can be alleviated by adopting a rule that shifts the burden of proof in some circumstances, as the EU proposal for an AI Liability Directive discussed in the next section does.

93 In his study of liability in multinational value chains Carston Koenig reaches a similar conclusion. Koenig explains that the main argument in favor of value chain liability is when some actors in the value chain are insufficiently receptive to liability incentives and, at the same time, there may be larger actors, such as multinational corporations, who as a function of their technical, legal, and economic relationship with the other actors in the value chain can influence their behavior. In such instances, it may be useful to direct liability at them. Koenig clarifies that the standard of liability should distinguish the type of actor: subsidiaries, independent business partners, and indirect business partners. Under regular subsidiary regimes, parent companies are held strictly liable for the actions of their subsidiaries. For business partners or indirect partners, Koenig suggests that leading companies be required to take "appropriate measures" to prevent and mitigate harm in the value chain. Thus, because the degree of influence is lower, the standard of liability should be lower and only fault-based. *See* Carsten Koenig, "An Economic Analysis of Value Chain Liability", *Harvard Law School*, 2024. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4819667.

The following table presents four examples of liability rules based on the principles discussed before:

TABLE 3		
AI LIABILITY ALONG THE VALUE CHAIN		
Rule	**Example**	**Liability rule application**
General rule: All actors in an AI value chain are under a fault-based regime and jointly several.	The AI-powered medical diagnosis system is developed by company A (the provider), integrated into its services by a hospital (the deployer) (B), and used by physicians (C).	The ideal situation is one when all actors take optimal care, that is, for example, the model developer conducts testing including for bias and edge cases; the hospital system does the integration, implements, thus rigorous validation and each hospital provides comprehensive training for doctors using the system; and doctors use it responsibly and do not trust the system blindly. All actors have incentives to take optimal measures of care because they may be held liable for the totality of harm if they don't. However, they will avoid liability if they can show that they took adequate measures of care

TABLE 3		
AI LIABILITY ALONG THE VALUE CHAIN		
Rule	**Example**	**Liability rule application**
Exception 1: AI products and services offered to non-professional natural persons: The case of product liability.	An AI developer puts an AI-enabled product on the market for consumers to use.	The producer is the cheapest cost avoider and consumers should not be expected to have expertise in handling AI-powered systems. Thus, the developer will be held strictly liable for a defect in the product regardless of whether it was his actions or the actions of its suppliers that caused it. However, he may avoid liability if the defect was "unknowable" at the time the harm occurred. The producer may still try to get compensation from guilty suppliers. The scope of liability will be limited if victims were contributory negligence, for example, if they did not comply with instructions.

TABLE 3		
AI LIABILITY ALONG THE VALUE CHAIN		
Rule	**Example**	**Liability rule application**
Exception 2: Quasi-strict liability for AI products and services offered by large market actors, with significant technical and economic control over its AI product or service and provider network. (This would be a product liability-like regime, but extendable to commercial transactions, too.)	A major technology service provider offers an AI assistant, developed by another party, for companies to incorporate in their different practices. A court finds that, in this case, the economic position and technical literacy of the technology company allows it to be in a position of control over the companies that provide it with software and components. This exception, however, should be used carefully, as it may affect incumbents. Large AI providers should only be held liable for the actions of their direct providers, and so long as the defect or fault was knowable at the time of occurrence.	The major technology company is the cheapest cost avoider and could be held strictly liable (as in a product liability regime) for all harm caused by a defective AI system or product for the actors in the supply chain it controls, irrespective of whether these harms originated in the actions of their providers or in their own. The producer may still try to obtain compensation from negligent suppliers. Given the system's design may only be in a position to take "best efforts" to ensure that downstream users use the system well. The firm that licenses the AI system and victim are however under a fault-based regime and must show that they took reasonable measures of care/acted according to the instructions of the system. Otherwise, the liability of the cheapest cost avoider may be diminished to the extent their actions contributed to harm.

TABLE 3		
AI LIABILITY ALONG THE VALUE CHAIN		
Rule	**Example**	**Liability rule application**
Exception 3: Exceedingly risky AI systems.	An AI system that manages critical infrastructure (such as defense mechanisms or the electrical grid of a city) is designed by a private company but managed by a government agency.[94]	Both actors are held strictly liable, and their liability is apportioned to the proportion to which they may have contributed to harm.[95]

Table 3: Examples of AI liability along the value chain

94 Different countries adopt different liability standards for critical infrastructure. For example, Germany typically uses a fault-based system, while Switzerland employs a strict liability system. Anne van Aaken and Isabelle Wildhaber, "State Liability and Critical Infrastructure: A Comparative and Functional Analysis", *European Journal of Risk Regulation*, 2015, Vol. 6, No. 2, pp. 244–254

95 As discussed above, without punitive damages this may still be a suboptimal rule, but a case specific analysis would be required.

05

The EU liability regime: A Discussion

05

The EU liability regime: A Discussion

In 2022, the European Commission proposed two Directives to address the issue of AI liability: an Artificial Intelligence Liability Directive (AILD) and a revision of the Product Liability Directive (PLD). The revised PLD was adopted and went into force in late 2024. The AILD, which from the begining had gathered less support, was withdrawn in February 2025.

This section presents how EU regulators were thinking about AI liability along the value chain, and offers some recommendations based on the framework proposed above for future attempts in the EU and elsewhere.

In general, the reforms passed and proposed in Europe are similar to the framework proposed in Section 4: (1) it defers to national law as the baseline liability regime, which in many cases will be a fault-liability regime, (2) it extends product liability for certain AI systems, consumer facing products, and includes certain provisions to ease the burden of proof of victims, and (3) it places most of its emphasis on alleviating the burden of proof of victims. As discussed in Section 4, this regime type has challenges.. Importantly, where product liability is not the applicable law, the regime will be rather fragmented where member states have different standards of protection for different activities (such as driving, or utility provision). Additionally, even though the alleviation of the burden of proof of victims is a good policy measure, establishing what are adequate standards of care will still be difficult for victims and courts alike.

As the liability directives are closely related to the framework created by the EU AI Act, the Section starts with an overview of the EU AI Act, and then briefly presents the PLD and the now withdrawn AILD. Readers familiar with these instruments can skip to the analysis starting in section 5.4.

5.1 The AI value chain in the AI Act

The cornerstone of EU AI regulation is the AI Act.[96] The AI Act is an umbrella and union-wide regulation that proposes a risk-based approach to AI regulation, which seeks to ensure that products and services integrate safety and security by design. It focuses on mandating that the actors that develop and place in the market AI systems that are considered to create heightened risks of harm comply with a series of risk-mitigation obligations. Subject to these obligations are mainly a list of so-called high-risk systems[97] and general-purpose AI models.[98] Risk mitigation obligations include implementing risk management,[99] furnishing technical documentation about the system's operation,[100] maintaining records on its functionality,[101] designing systems to facilitate user comprehension and human oversight,[102] amongst others.

The AI Act captures the AI value chain by focusing on two main types of actors: AI providers and AI deployers.

AI providers: The bulk of the obligations under the AI Act fall upon so-called AI providers. AI providers are the natural or legal persons who develop an AI system, have it developed, place it on the market, or put it into service under its name or trademark.[103] Providers are the ones who must ensure that their high-risk systems comply with all the obligations of the Act,[104] and must keep documentation, have a quality management system, and ensure that their systems conform with the regulation in general.[105] Providers of general-purpose AI models must also draw up and keep up-to-date technical documentation[106] and if the model presents systemic risks perform model evaluations and adopt mitigation measures.[107]

96 European Commission, "Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things, and Robotics", *Brussels*, February 19, 2020, COM(2020) 64 final., at 4.

97 *See* AI Act (n21), Article 6. As listed in Annex III These include systems that are either intended to be used as a safety component in motor vehicle security, those used in the management and operation of critical infrastructures, like road traffic or the supply of utilities, biometric identification systems, and AI systems intended to be used in educational and employment settings to determine, respectively, access to institutions or recruitment

98 AI Act (n21) Article 51, 52, 53

99 AI Act (n21) Article 9

100 AI Act (n21) Article 11

101 AI Act (n21) Article 12

102 AI Act (n21) Article 14

103 AI Act (n21) Article 3(3)

104 AI Act (n21) Article 16 (a)

105 AI Act (n21) Article 16.

106 AI Act (n21) Article 53

107 AI Act (n21) Article 55

AI deployers: AI deployers are the natural or legal persons who use an AI system under their authority "except where the AI system is used in the course of a personal non-professional activity."[108] The obligations of deployers of high-risk systems are to take appropriate measures to ensure that they can be used per the instructions given by the provider, assign human oversight where necessary and monitor its operations.[109] In some instances, they also must conduct a fundamental rights impact assessment.[110]

Deployers have more substantive obligations the more control they have over the system. If they have control over the input data, then they are also responsible for ensuring that it is relevant for its intended purpose.[111] Similarly, deployers that put their name or trademark on a high-risk system or a general-purpose system, modify it, or modify its intended purpose, are considered the provider, and have the same obligations.[112]

5.2 The Revised Product Liability Directive

The Revision of the Product Liability Directive (PLD) was adopted in late 2024. It replaced the PLD adopted in 1985. This new version seeks to adapt the EU's product liability regime to new technologies — such as software and AI systems — and addresses some of the difficulties injured people face in complex cases to prove that a product was defective and caused the damage they suffered.[113]

To do so, the PLD establishes a principle of strict liability for the manufacturers of a product or a subsequent actor that modifies it, for damages suffered by natural persons and caused by defective products.[114]

Products are defined as any tangible or intangible movable good, including software.[115] A defective product is defined as one that "does not provide the safety that a person

108 AI Act (n21) Article 3(4)

109 AI Act (n21) Art 26.1, 26.2, 26.5

110 AI Act (n21) Article 27

111 AI Act (n21) Article 26.4

112 AI Act (n21) Article 25.

113 *See* European Commission, Explanatory Memorandum of the Proposal for a Directive of the European Parliament and of the Council on liability for defective products Brussels, 28.9.2022 COM(2022) 495 final 2022/0302(COD) explaining that "the rules excessively limited the possibility of making compensation claims. Property damage worth less than Eur 500 is not recoverable under the PLD.

114 Directive (EU) 2024/2853 of the European Parliament and of the Council of October 23, 2024 on liability for defective products and repealing Council Directive 85/374/EEC (Text with EEA relevance), PE/7/2024/REV/1 [*hereinafter* PLD] Article 1, Recital 9. *See also* Article 8 explaining that economic operator refers mostly to manufacturers and those that modify products.

115 PLD (n115) Article 3, Recital 13

is entitled to expect" or is required by law.[116] The assessment of the defectiveness of a product is done based on criteria such as the technical features of the product, the packaging, its reasonably foreseeable use, and the moment in time when it was placed on the market or where it leaves the control of the manufacturer.[117]

Economic operators are actors that can be held liable are manufacturers, providers, authorized representatives and importers.[118] From a value chain perspective, note that this means that injured persons can seek compensation from both the manufacturers of the component and the manufacturer of the product.[119] They will be jointly and severally liable.[120]

Damages are mainly personal damages (including death) and damage to property or data.[121] Deployers of free and open software that use it as a component of a product that is then placed on the market in the course of a commercial activity may be held liable for the defect in such software.[122]

Burden of proof: The PLD maintains that it is on the claimant to prove the elements of product-strict liability: defectiveness of the product, the damage suffered, and the causal link between the effectiveness and the damage.[123] However, it includes a series of measures to alleviate the information asymmetries that may often exist between victims and economic operators. First, claimants who have presented enough facts and evidence to support the plausibility of their claim are entitled to request from the defendant evidence at their disposal.[124] Second, the defectiveness of the product will be presumed if the defendant fails to disclose relevant information, if the claimant demonstrates that the product does not comply with mandatory safety requirements, or if the claimant shows the damages were caused by an obvious malfunction.[125] Third, the causal link between the defectiveness of the product and the damage will be presumed where the defectiveness has been established and the damage "is of a kind typically consistent with the defect in question." [126]

116 PLD (n115) Article 7.1
117 PLD (n115) Article 7
118 PLD (n115) Art 4(15); Art 8
119 PLD (n115) Article 8, Article 12
120 PLD (n115) Article 12
121 PLD (n115) Article 6
122 PLD (n115) Recital 15
123 PLD (n115) Article 10
124 PLD (n115) Article 9
125 PLD (n115) Article 10.2
126 PLD (n115) Article 10.3

Economic operators will not be liable, however, if they show that it is probable that the defectiveness that caused the damage did not exist at the time it was placed in the market, or the scientific and technical knowledge at the time the product was placed on the market was not enough to discover the defect.[127] Liability may also be reduced or disallowed if the damage was caused by the fault of the injured person.[128]

5.3 The Artificial Intelligence Directive (AILD)

The AILD was a proposed directive that layed down common rules to mainly address the challenges victims face in identifying who is liable and proving their claims, which can be difficult or costly due to the nature of AI systems.[129] The AILD was withdrawn by the European Commission in early 2025, primarily due to mounting opposition, overlapping concerns with the Product Liability Directive, and potentially broader political shift toward deregulation in Brussels. The question of liability in the AI value chain remains crucial, however, because the current framework doesn't fully address how responsibility should be distributed along the value chain, for example, when damage is caused by defects in underlying AI models, potentially leaving downstream players vulnerable for issues originating at the model level.[130]

The directive followed a "minimum harmonisation approach" and, consequently, though it assumed that the baseline standard is fault-based liability, member states can maintain their existing liability. Thus, national laws could maintain other provisions reversing the burden of proof under national fault-based regimes, or even regimes with a higher standard, such as strict liability.[131]

The proposed directive contained two main measures: First, national courts were empowered to demand the disclosure of relevant evidence from high-risk systems suspected of having caused damage to providers or those subject to their obligations when certain conditions are met. If defendants did not disclose information, there would be a presumption of fault.[132]

127 PLD (n115) Article 11

128 PLD (n115) Article 13.2

129 European Commission, Explanatory Memorandum, Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), COM/2022/496 final, article 1 [*hereinafter* Proposal AILD].

130 *See* Luca Bertuzzi, *Twitter thread* February 12, 2025. Available at: https://x.com/BertuzLuca/status/1889713305863786924

131 Proposal AILD (n130) Recital 14

132 Proposal AILD (n130) Article 3

Second, the withdrawn directive established that courts will presume the causal link between the fault and the output where the claimant (1) had established the existence of harm and the fault of the defendant; and (2) it was "reasonably likely, based on the circumstances of the case," that the fault or the failure of the AI system led to the output, and (3) the claimant demonstrated that the output or the failure of the AI system gave rise to damage.[133]

Additionally, the AILD established that fault would also be presumed where, in the case of high-risk systems as defined by the AI Act, the complainant had demonstrated that the provider or the deployer failed to comply with any of the requirements and obligations under the AI Act.[134]

In the case of non-high-risk AI systems, courts could determine that it is excessively difficult for the claimant to prove the causal link between damage and fault and presume it.[135]

133 Proposal AILD (n130) Article 4.1

134 Proposal AILD (n130) Article 4.2

135 Proposal AILD (n130) Article 4.5

TABLE 4	
KEY DIFFERENCES BETWEEN THE PLD AND THE AILD	
PLD	**AILD PROPOSAL**
Claim rooted in EU law	Claim rooted in Member State law
Material and procedural aspects of product liability	Procedural aspects of non-contractual civil liability for AI systems
Applicable to physical products and software, including AI systems	Applicable to AI systems only
Supposedly strict liability	Fault-based liability (or following member state law)
Claims against manufacturers and other entities in the supply chain	Claims against manufacturers, professional users and consumers
Eligible damage: privately used property, death or personal injury, and data loss	Eligible damage: potentially also professionally used property, fundamental rights and primary financial loss
Full harmonisation	Minimum harmonisation

Table 4: Key Differences between the PLD and the AILD

Source: P. Hacker, 'The European AI Liability Directives - Critique of a Half-Hearted Approach and Lessons for the Future', Computer Law & Security Review, Vol. 51, Article 105871, 2023, p. 7.

5.4 Analysis of the EU Approach to AI Liability

This section analyzes the AI approach to AI liability mainly based on the framework summarized in Section 4.4.[136]

In general, the AILD and, in particular, the PLD were similar to the framework proposed in Section 4. EU regulators considered as a baseline a fault-based liability regime or whatever regime is in place in national liability laws, and increased the standard by extending product liability to certain AI systems, consumer facing products, and includes certain provisions to ease the burden of proof of victims.

As discussed in Section 4, however, that regime does not come without its challenges, which aided in the withdrawal of the AILD. Further, given that the AILD is a minimum harmonization directive, the AI liability landscape would remain somewhat fragmented. More substantively, proving fault or a defect is, and was going to remain, complicated in cases involving complex AI systems. This Report argues that regulators can address this by further empowering courts to order disclosures of evidence for complex systems, or shifting the burden of proof. It does not suggest a complete overhaul in favour of strict liability, given that many systems in Europe will soon be subject to a new regulatory burden under the AI act.

What follows analyzes the EU's approach to AI liability focusing on these key policy questions, as developed in Section 4: 1) The chosen liability standard, 2) how liability is shared among multiple tortfeasors, and 3) alleviating information asymmetries.

a) The choice for liability regime

The AILD maintained member state regimes, where the general rule is fault-based liability even if it allowed them to extend strict liability that may apply to damage caused by AI systems or activities governed by strict liability standards.[137] Importantly, the EU effectively extended product liability for products placed on the market for natural persons to use, there is a product liability regime. This latter choice aligns with the direction of the general framework proposed in Section 4, and it highlights EU regulators' interest in avoiding disincentivizing desirable innovation with a too-strict liability regime.

136 Other notable analysis are Hacker, *Impact Assessment* (n67), Sandra Wachter, "Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond", *Yale Journal of Law & Technology*, Vol. 26, Issue 3, July 1, 2024.

137 Proposal AILD (n130) Recital 4

As discussed at length in Section 4, a fault-based liability regime has the great advantage of being more fair as it only holds responsible AI actors in the value chain that fail to take adequate measures of care. Thus, for example, in the case of an accident involving an automated or a semi-automated vehicle, the different actors involved — such as the driver, the car manufacturer, and the navigation system developer — will all be able to exempt themselves from liability if they show that they took adequate measures of care).

As the example already suggests, the main challenge with a fault-based liability regime is how we determine that adequate safety measures have been taken (and that more should not have been taken). This is not an easy question to answer, as both victims and courts may struggle establishing what fault entails. Note that the AILD did not solve this problem, because it did not change underlying national liability regimes and didn't, by itself, address the question of what "fault" or reasonable care meant.

In the PLD, this same problem persists. As discussed in Section 4.1, product liability is not a full strict liability regime and proving that a defect, a deviation from the standards that are reasonable to expect at the time the product leff their control — will remain hard. It is worth noting, however, that the inclusion of "the moment in time when the product was placed on the market or (...) the moment in time when the product left the control of the manufacturer" is a very interesting inclusion into the revised PLD, as it acknowledges that the actors of the AI value chain may maintain significant control over an AI system even after they have licensed or sold it — such as car manufacturers that maintain control over the navigation system or other elements of the car after it has left their stores. This inclusion recognizes, and highlights, that the responsibilities of different actors of the AI value chain coincide and overlap in time in several circumstances.[138]

Specifically in the EU context, where the AI Act has been enacted, one implication both for any future liability regulation, but still for courts who facing AI liability claims under existing liability rules is that courts and AI actors will rely on the compliance of these obligations to show that they took "due measures of care," especially for high-risk systems or AI otherwise subject to risk-mitigation obligations, such as general purpose AI models. Indeed, the explanatory memorandum of the AILD explained that the fault of the defendant has to be proven according to the applicable Union or national laws: "Such fault can be established, for example, for non-compliance with a duty of care pursuant to the AI Act or pursuant to other rules".[139] Additionally, since compliance with the obligations of the AI Act will be easiest shown by demonstrating conformity with harmonized standards, it is not unlikely that the standards of care under any liability regulation will

138 PLD (n115) Article 7.2(e)

139 Proposal AILD (n130), Explanatory Memorandum, Presumption of Causal Link Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0496, Proposal AILD (n130) Article 4

bend towards compliance with these technical standards.[140] This occurs in a similar way under the PLD, where Article 11 exempts economic providers from liability economically in cases where the defect "that caused the damage is due to compliance of the product with legal requirements."[141]

The challenge with the assimilation of standards of care with regulatory technical standards may, as discussed in Section 4.1, discourage some AI actors, particularly the most sophisticated ones who have important expertise, to choose the preventive measures that are required by law or typically required by courts, instead of those they believe are most effective. This can diminish the usefulness of their overall harm-mitigation strategy, especially as systems evolve and AI actors could be quicker at adopting mitigation strategies than regulators at adopting regulations. It could be argued that a way to avoid this is to extend strict liability to high risk systems. In the EU, however, where regulation has been adopted, requiring compliance with regulatory standards and holding providers to strict liability may be too burdensome at least in certain instances.

At the same time, it may be not entirely undesirable that regulation does guide the harm-mitigation strategy of less sophisticated actors in the AI value chain.[142] Similarly, since there are only a limited number of AI systems that are subject to regulation under the AI Act, assimilation of standards of care with regulatory standards only applies to a limited set of AI systems that are subject to significant regulations under the AI Act.

Because of the challenges of fault based liability, commentators like Philip Hacker have thus proposed that it would be perhaps cheaper, and even provide more legal certainty, to adopt a strict liability regime.[143] There could, indeed, be potential for EU regulators or even member states to consider extending a strict liability regime for very specific scenarios of damage by AI systems (under some national liability regimes, this may be already the case). Further research should be conducted on the potential effects of this heightened liability burden.

When considering the complexity and the multiplicity of actors involved in the AI value chain, however, the choice for a liability standard should not be discussed independently from how these multiple actors will divide liability amongst them, as this policy choice

140 Beatriz Botero Arcila, "AI Liability in Europe: How Does It Complement Risk Regulation and Deal with the Problem of Human Oversight?", *Computer Law & Security Review*, Vol. 54, September 2024, 106012, at 16, also Hacker, *Impact Assessment* (n67)

141 PLD (n115) Art 11(d)

142 It is unclear, however, how this may play out. Elsewhere I argued that the success of liability claims strongly depends on compliance with AI Act standards. High-risk systems are more likely to meet and have standards focused on guiding regulatory fitness. Non-high-risk systems may lack detailed documentation to support such claims, creating potential disparities in legal protection across different AI system categories. *See* Botero Arcila (n141)

143 Hacker, *Impact Assessment* (n67)

interplays with liability regimes affecting their incentives (see Section 5.2), this is thus what we do next.

b) How liability is shared amongst multiple tortfeasors

The EU AI liability framework only partially addresses directly how liability is shared amongst multiple tortfeasors. The AILD did not address this question because it deferred to national liability systems specifically. This consequently varies and will continue to vary from country to country. The framework above strongly favoured a joint and several liability regime, which is the one mostly adopted by the PLD:

The PLD establishes joint and several liability for instances where two or more economic operators, as defined by the PLD, are liable for the same damage.[144] This can be the case, for example, of a defective component integrated into a product that causes damage. As recital 53 explains, in such circumstances "the injured person should be able to seek compensation both from the manufacturer that integrated the defective component into its product and from the manufacturer of the defective component itself."[145]

The combination of joint and several liability with product liability offers several benefits. In particular, it recognizes the often overlapping nature of the roles of different actors in the AI value chain, it encourages risk-mitigation strategies amongst the different economic operators. At the same time, it recognizes that in many circumstances economic operators will have the capacity to enforce quality standards on each other and allocate responsibility to certain providers via private ordering. Let's look at each of this in more detail:

Article 8 of the PLD establishes that manufacturers of a defective product and manufacturers of defective components, "where that component was integrated into, or inter-connected with, a product within the manufacturer's control and caused that product to be defective" can be held jointly and severally liable for harm.[146] At the same time, the PLD positions the product manufacturer as a cheapest cost avoider where the component is "integrated into, or inter-connected with, a product within the manufacturer's control."[147]

144 PLD (n115) Article 12.1

145 PLD (n115) Recital 53

146 PLD (n115) art 8.1

147 *See* PLD (n115) Article 8.1, last subparagraph

This arrangement manages to create important incentives for component manufacturers because under joint and several liability they may ultimately be held liable for more damage than their defective product caused. They may avoid full liability, however, if they can show that their component meets the standard of "the safety that a person is entitled to expect."[148] At the same time, this arrangement positions the product manufacturer where the component is "integrated into, or inter-connected with, a product within the manufacturer's control,"[149] as a cheaper cost avoider, while explicitly extending joint and several liability for the defect in the component. This creates extra incentives for the product manufacturer to supervise the actors in the supply chain, and allocate clear responsibilities to them under private agreements. In doing so, the PLD does a good job at preventing the raise to the bottom described in Section 4.1.[150]

A vacuum left by the rule, as written, is the case for damage caused by components when the product is no longer under the control of the main product manufacturer. It may be that product manufacturers and component manufacturers address this type of instance via their own contractual agreements, but it may be that a potential victim may have to rightfully identify the manufacturer of the component to seek relief for harm, which may prove exceedingly hard.

Lastly, it is worthy to note and praise that, again the PLD assigns liability based on the control any given economic operator may have over the AI systems and its outputs and thus, extends the role of manufacturer for any person "that substantially modifies a product outside the manufacturer's control and thereafter makes it available on the market or puts it into service."[151]

c) Access to evidence, the burden of proof, and information asymmetries

One of the main focuses of the EU AI Liability discussion is to facilitate access to evidence and facilitate the proof of certain elements of liability for victims of harm. As described in section 5.3, this was the objective of the two main measures of the AILD, and the PLD includes a provision that mandates the disclosure of evidence, when the

148 *See supra* Section 4.1

149 *See* PLD (n115) Article 8.1, last subparagraph

150 The point was that In a strict liability regime with joint liability, actors in the AI value chain might be incentivized to reduce their care measures. If one actor believes another isn't preventing potential harm effectively, they may lower their own standards. This creates a "race to the bottom" where actors prioritize avoiding potential costs over preventing harm, leading to a problematic free rider situation. *See* section 4.1

151 PLD (n115) Article 8.2

defendant has presented "facts and evidence sufficient to support the plausibility of the claim for compensation."[152]

This is consistent, in general, with the observation that AI systems' opacity and complexity complicate victims' chances of success in liability procedures, and are well-intentioned measures to address this information asymmetry.

At the same time, these provisions strongly rely on the assumption that such evidence will exist and, importantly, that victims and courts will be able to understand it. Reading the PLD together with the AI Act, suggested that some of these evidence disclosure obligations may be limited, as only high-risk AI systems are required to maintain comprehensive documentation, potentially leaving victims of non-high-risk systems with less access to critical evidence.[153] An important limitation of the AILD, was that evidence disclosure measures applied mostly to providers of high-risk systems.

In future AI liability initiatives it would be desirable to extend the provision already included in the PLD, which empowers courts to order the disclosure of evidence for providers of complex AI systems, when the person seeking compensation "has presented facts and evidence sufficient to support the plausibility of the claim for compensation" and has also duly justified why it is difficult to prove the elements of liability without having access to such evidence.[154] This would better allow victims of harm to pursue their claims.

152 PLD (n115) Article 9

153 *See* making this point Botero Arcila (n141); *see also* Proposal AILD (n130) Art 3.1.

154 *See* PLD (n115) Article 9.1

Conclusion

AI raises new concerns and particular challenges for civil liability and for the ability of individuals who were harmed by AI systems to seek redress. At the same time, different liability regimes create varying incentive structures and costs for AI developers and other actors across the complex AI value chain. As regulators around the world are grappling with these questions, this Report sought to provide an overview of the dynamics at play here and proposes a preliminary framework for adequately addressing them. The framework proposes as a baseline rule a fault-based, joint and several liability regime to balance the interests of (potential) claimants and actors along the AI value chain. It also proposes extending strict liability or product liability on a case-by-case basis to address some of the baseline rule's limitations.

The PLD is aligned with the framework proposed in Section 4, to the extent that it enhances access to evidence by victims of harm and, importantly, extends product liability to consumer products that involve AI systems. Importantly, joint and several liability for AI providers and components is an important innovation that acknowledges the complexity of the AI value chain, strengthens providers and components' suppliers incentives to take care and may facilitate victim's compensation.

It is unclear, at the time of writing, whether the EU will try to propose new rules for AI liability in general. Worth rescuing from the AILD, however, are the measures that would have facilitated access to information and shifted the burden of proof of causation, as they were rightly intending to facilitate victims' access to compensation without overburdening AI actors.

Further, as European regulators, and regulators around the world, discuss potential reforms to liability regimes it would be highly desirable to conduct further empirical research on the costs of liability regimes and how they relate with regulatory compliance and the ease with which victims of harm involving AI systems can pursue their claims. This would greatly improve the policy debate and our understanding of how liability along AI value chains is distributed and enforced through private ordering. It would provide valuable insights into the current practices and challenges faced by victims of harm, as well as by the different actors in the AI value chain. By examining these private arrangements, regulators can better identify the gaps and inefficiencies in the existing frameworks. This understanding will help shape more effective and targeted interventions, ensuring that liability is appropriately managed and that the regulatory environment supports innovation while protecting the rights and interests of all parties involved.

Similarly, regulators and policymakers considering civil liability, in general, should consider whether the administrative and social costs of litigation and the liability system are substantial, and whether victims will indeed bring these claims.[155] In the EU, policymakers should further consider the costs and benefits of harmonizing its liability regime for AI vis-à-vis the more fragmented status quo across member states.

Considering these various factors will allow policymakers to better assess the various trade-offs involved in developing (or not developing) dedicated AI liability regimes, and to balance innovation, justice concerns and legal certainty for business and harmed individuals alike.

155 *See* Shavell, Liability for Accidents (n62) at 16

Annex

Annex I. Private ordering and AI liability along the value chain

This Annex develops further and briefly presents the way different AI actors govern AI liability, absent regulation, through their terms and conditions, licenses, and contracts regulate the obligations and responsibilities of the different actors of an AI supply chain. This is what is meant by "private ordering." Contracts amongst AI actors often allocate to specific actors' liability for specific damages — such as for outputs that infringe copyright[156] — or to specific amounts.

These instruments are varied, however, and a comprehensive report on the nature and trend of these instruments goes beyond the scope of this Report.

1.1 Generative AI terms of use and licenses

Some scholars argue that generative AI companies tend to defer liability for outputs and content downstream in the value chain, to users or providers. This "platformisation paradigm," resembles social media policy and regulation that similarly tends to exempt platforms from liability for the information they carry.[157]

OpenAI's terms of use, for example, include a "Disclaimer of warranties" which explains that their services are provided "as is." That is, except to the extent prohibited by law, OpenAI, affiliates, and licensors make no promises or guarantees, express or implied, regarding the condition of their product. This no-warranty includes the following:

> "merchantability, fitness for a particular purpose, satisfactory quality, non-infringement (...) we do not warrant that the services will be uninterrupted, accurate or error-free, or that the content will be secure or not lost or altered.
>
> You accept and agree that any use of outputs from our services is at your sole risk and you will not rely on outputs as a sole source of truth or factual information, or as a substitute for professional advice."[158]

156 Edwards (n40) at 13

157 Edwards (n40) at 21

158 OpenAI (n42)

Regarding the user, the terms of service establish that

> "Neither we nor any of our affiliates or licensors will be liable for any indirect, incidental, special, consequential, or exemplary damages, including damages for loss of profits, goodwill, use, or data or other losses, even if we have been advised of the possibility of such damages."[159]

Similarly, referring to downstream providers or corporate AI users, OpenAI's terms of service establish that,

> "...if you are a business or organization, to the extent permitted by law, you will indemnify and hold harmless us, our affiliates, and our personnel, from and against any costs, losses, liabilities, and expenses (including attorneys' fees) from third party claims arising out of or relating to your use of the Services and Content or any violation of these Terms."[160]

Llama's 3.2 Community License Agreement, also grants "a non-exclusive, worldwide non-transferable and royalty-free limited license under Meta's intellectual property or other rights owned by Meta embodied in the Llama Materials to use, reproduce, distribute, copy, create derivative works of, and make modifications to the Llama Materials."[161] The license includes a variety of obligations for users or providers relying on Llama — such as displaying "Built with Llama" on a related website to their products and complying with applicable laws.[162]

At the same time, the license includes a very similar disclaimer of warranty which establishes that "Llama materials and any output and result... are provided on an "as is" basis... in no event will Meta or its affiliates be liable under any theory of liability."[163] In sum, users (or providers using Meta) use Llama at their own risk.

159 *Id.*

160 *Id.*

161 Meta, Llama 3.2 Community License Agreement 1.a https://raw.githubusercontent.com/meta-llama/llama-models/refs/heads/main/models/llama3_2/LICENSE

162 Meta (n162) 1.b.i; 1.b.iii

163 Meta (n162) 3, 4.

1.2 B2B AI terms of service

Not all terms of service displace liability fully to users, however. The terms of service of companies offering more specific AI services and solutions to businesses establish obligations amongst the parties and limit liability to the fulfillment of those obligations. In such instances, thus, they allocate liability a little more evenly amongst developers and users. How these agreements work, however, is harder to understand as they are not often made public.

MicroAI, for example, is a company offering AI and machine learning products for machinery and network optimization, predictive maintenance, and cyber security. The warranty clause of its terms of service allocates more responsibility on itself and warrants that the service it offers will operate "in substantial conformity with the applicable Documentation,"[164] by which it means its description of its services and products.

The warranty, however, does not apply "if the error or non-conformance was caused by misuse of the Service or Deliverables, modifications to the Service or Deliverables by Customer, a User or any other third-party, or third-party hardware, software, or services used in connection with the Service."[165]

To give another example, Amazon Web Services' (AWS) Terms of Use establish that AI Services are not intended for use in critical systems or hazardous environments and are not intended to be used for clinical decision-making. Users "are responsible for liability that may arise in connection with any such uses."[166]

However, and as some other companies have done, AWS's Terms of Use also establish that AWS will defend users and your team against third-party claims that Generative AI Outputs from their services infringe on intellectual property rights and will pay any resulting judgments or settlements. However, they exempt themselves from being liable in several circumstances such as if the infringement is due to users' inputs, ignoring their instructions, breaching the Terms of Use, modifying the AI services, or continuing to use the output after being told to stop.[167]

Interestingly, to the extent that some of AWS' services use models developed by third parties, such as Llama or Mistral, AWS Terms of Use include extra conditions, which often refer or copy to that provider's licenses or terms of service.[168]

164 Micro AI, Terms of Service, Version Number 09062022 https://micro.ai/terms-of-service

165 Micro AI (n165)

166 Amazon, AWS terms of service, at 50.6. (November 7, 2024 version)

167 Amazon (n167)

168 Amazon (n167) 50.12.5. and 50.12.6

1.3 Suppliers of components: The case of LAION

Lastly, AI developers rely on suppliers of components — from hardware to data to labor — who may also be relevant for AI outcomes. We look here at a data provider, as data inputs are very important for AI outputs.

The Large-scale Artificial Intelligence Open Network (LAION) is a non-profit organization that "provides datasets, tools, and models to liberate machine learning research."[169] LAION is one of the largest providers of publicly available data, and it has been used by start-ups and academics to develop AI and, possibly, by large tech companies too. It has also been under the public eye because its data sets have included copyrighted material and even child sexual abuse material.[170]

LAION is a non-profit organization, which indexes to the internet so it does not host actual image data. It releases open datasets indexed to the source on the internet and it does so under the most open Creative Commons license that only requires attribution.[171] The license available on the GitHub page establishes that the software is provided "as is" without warranty of any kind. Additionally, because it is a research organization they are covered under an EU law exemption that allows research organizations to use copyrighted materials to research AI.[172] Thus, LAION suggests users of their reconstructed databases clean the data from potentially copyright infringing or disturbing data.[173]

1.4 Conclusion: The relationship of private ordering and non-contractual civil liability for AI systems

The different actors of AI value chains often distribute amongst them different responsibilities, and the burden of liability via private ordering agreements and documents, such as terms of service, licenses and contracts. These are important institutions because even if victims of harm can bring a claim against all the actors that potentially caused it regardless of whether they have a contractual relationship with them, courts will look at

169 LAION https://laion.ai/ (Last visited January 28, 2025).

170 David Thiel, "Investigation Finds AI Image Generation Models Trained on Child Abuse," The Cyber Policy Stanford University, December 20, 2023. https://cyber.fsi.stanford.edu/news/investigation-finds-ai-image-generation-models-trained-child-abuse

171 Christoph Schuhmann, "LAION-400-Million Open Dataset," August 20, 2021. https://laion.ai/blog/laion-400-open-dataset/

172 Andrés Guadamuz, "LAION Wins Copyright Infringement Lawsuit in German Court," TechnoLlama, September 28, 2024 https://www.technollama.co.uk/laion-wins-copyright-infringement-lawsuit-in-german-court

173 LAION, FAQ https://laion.ai/faq/

these arrangements to understand the expected role of the different actors.[174] They will also uphold the contractual provisions agreed upon by the different parties of the AI value chain unless these are considered to go against the law and public order.

Thus, for example, in an accident involving a semi-autonomous vehicle, the victim of harm could bring a suit against the driver, the vehicle's manufacturer and the provider of the navigation system. (Compare this with an equivalent liability claim in an accident involving a non-automated vehicle, where the natural person to sue would be just the driver.) To be able to bring a claim against the vehicle manufacturer and the navigation system provider, the victim would need to be able to show, however, that it is at least likely that the vehicle's manufacturer or the navigation system developer, also contributed to harm.

For the reasons explained in Section 2, the victim of harm may face important difficulties in understanding a priori who the participants of a given AI system's value chain are and who could be potentially liable for the damage caused.[175] This is something that, for example, the European AI Liability Directive seeks to address, as will be discussed in Section 7.

In this example of autonomous vehicles, the victim of harm has no contractual relationship with any of the actors, so this would be a traditional liability claim. However, the vehicle manufacturer, the navigation system provider, and potentially the driver may have contractual relationships amongst them, however.

What follows in a liability procedure is thus examining these contractual relationships to establish who should continue in the process. If they are found to conform with existing law, they will usually be respected. This is unless the victims of harm can show evidence of fault, a product defect, or the breach of another applicable legal obligation, depending on the jurisdiction.

Thus, for example, something that would have to be determined in court is the validity of a clause in the terms of service where the vehicle manufacturer may have explicitly warned drivers to keep their hands on the wheel and be "prepared to take over at any moment." The plaintiff would also have to prove that they complied with their side of that

174 This wasn't always the case. In the United States, for example, the famous case that first allowed this was *MacPherson v. Buick Motor Co*, in 1916. Previously the general rule had been that only the parties of a contractual relationship with a product's manufacturer could sue for the products malfunctioning and the damages they caused. *See* Marler Clark, "An Introduction to Product Liability Law," https://marlerclark.com/pdfs/intro-product-liability-law.pdf .

175 *See* Proposal AILD (n130) Recital 17 "The large number of people usually involved in the design, development, deployment and operation of high-risk AI systems, makes it difficult for injured persons to identify the person potentially liable for damage caused and to prove the conditions for a claim for damages."

deal.[176] Similarly, the navigation system provider may have warned users that the system is provided on an "as is" basis and that drivers are responsible for liability caused in connection with its outputs. If the navigation system is fully embedded in the vehicle, however, the contract between the navigation system developer and the vehicle manufacturer may, as in the examples above, limit the navigation system's provider liability. A court will have to determine whether this limitation is lawful.

At the end of this initial stage, the court will decide whether these contractual terms are enforceable and whether they apply to the situation and establish who will be a party to the liability claim. In the example above, imagine that the contract between the navigation system provider and the vehicle manufacturer is enforceable. However, the court may want to consider the liability of the vehicle manufacturer, as well as the one of the drivers, and so the case will continue with the car manufacturer and the driver as defendants.[177]

176 Mike Spector, Dan Levine & Mike Spector, "Exclusive: Tesla Faces U.S. Criminal Probe over Self-Driving Claims," Reu ters, Oct. 27, 2022, https://www.reuters.com/legal/exclusive-tesla-faces-us-criminal-probe-over-self-driving-claims-sources-2022-10-26/ (last visited Aug 25, 2023).

177 See *Hamida Begum v Maran (UK) Limited* [2021] EWCA Civ 326 (where the UK Court of Appeal allowed a negligence claim to proceed regarding a UK-based shipping agent whose sold oil tanker led to a shipbreaker's death in Bangladesh. Despite doubts, the court ruled the duty of care claims were substantial. This highlights the growing risk of corporate liability for UK companies in their global value chains, reflecting the English Courts' increasing willingness to explore such claims. Available at: https://www.bailii.org/ew/cases/EWCA/Civ/2021/326.html

Authors and Acknowledgments

This paper is written by Beatriz Botero Arcila, Assistant Professor of Law, Sciences Po Paris Law School and Faculty Associate at the Berkman Klein Center for Internet and Society at Harvard University.

The author wishes to thank Giovanna Hajdu Hungria da Custódia for her research assistance and Natalia Galvis Arias, the participants of the Transatlantic AI and Law Institute Paris convening in October 2025, Mina Mohammadi, Maximilian Gahntz, Stefan Baack, Claire Jenifer Perhan, and Steven Shavell for their valuable feedback on this work.

This document was designed by Spitting Image, Bengaluru.