

November 2022

# FLI Position Paper on AI Liability

FLI position on the proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)

## Executive summary

*The Future of Life Institute (FLI) welcomes the opportunity to provide feedback to the European Commission on its proposal for an artificial intelligence (AI) Liability Directive<sup>1</sup>. Liability is an important instrument for safeguarding the interests of society. It can play a role in catalysing innovation by encouraging organisations to develop risk-mitigating technologies that reduce the likelihood of harm in products and services. At the same time, it prioritises the rights of individuals and can lead to increased trust and uptake in new technologies.*

*FLI welcomes the adoption of the new proposal to update liability rules to the digital age in a way which protects the rights of those harmed by AI systems and which encourages AI developers to make their products safer. To ensure legal clarity and a proper level of protection for victims harmed by AI systems across the EU, FLI makes the following recommendations: i) introduce a strict liability regime for high-risk and General Purpose AI systems; ii) establish a fault-based liability regime with a reversed burden of proof for other AI systems; iii) include a definition of damages that extends to immaterial and societal harms; and iv) establish clear rules on AI liability across the value chain.*

---

<sup>1</sup> Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence, COM(2022) 496 final, 28.9.2022.

## Introduction

The Future of Life Institute (FLI) is an independent non-profit organisation that works on maximising the benefits of technology and mitigating its risks. We created one of the earliest and most influential sets of AI governance tools - the Asilomar AI principles - and maintain a large network among the world's top AI researchers. In addition, we are the civil society champion of the recommendations on AI in the UN Secretary General's Digital Cooperation Roadmap, alongside the governments of France and Finland. FLI expanded its operations to Europe in 2021 and is registered in the EU Transparency Register under 787064543128-10.

FLI welcomes the adoption of the new proposal to update liability rules to the digital age. The regulation of liability has key social and economic functions.<sup>2</sup> In particular, liability rules can dissuade actors from engaging in risky activities,<sup>3</sup> thereby promoting safety standards that prevent and reduce accidents. They can also facilitate the correct pricing of a product or service and encourage innovation investment by mitigating uncertainty over the litigation process.<sup>4</sup> Furthermore, liability rules that distribute risks between market actors and establish mechanisms for compensation of damages encourage consumer uptake of technology and promote justice, equity, and fairness.<sup>5</sup>

The proposal for an AI liability Directive (AILD) applies to non-contractual fault-based civil law claims for damages caused by AI systems. It lays down rules that ease the burden of proof in relation to two aspects: disclosures of evidence and rebuttable presumptions of causality.

In particular, the AILD enables courts to order the disclosure of relevant evidence about specific high-risk AI systems suspected of having caused damage. In case of non-compliance with such an order, the AILD introduces a rebuttable presumption of non-compliance with a duty of care. Secondly, the AILD alleviates the burden for claimants to establish a causal link between non-compliance with a duty of care and the damage caused by introducing rebuttable presumptions of causality under very limited and complex conditions.

These rules are a welcome attempt to address AI liability at the EU level. Their enactment has the potential to close some liability gaps of the product liability proposal (PLD),<sup>6</sup> such as offering protection to both natural and legal persons harmed by AI systems and allowing for a broad scope of damages to be covered. However, as long as the AILD fails to provide for more legal clarity and the same level of protection across the EU, it will be insufficient to safeguard victims of AI systems.

As acknowledged by the European Commission, victims should enjoy the same level of protection for damage caused by AI as by other technologies.<sup>7</sup> FLI fully agrees and makes the following recommendations to ensure proper protection for damages caused by AI systems:

- I. Strict liability for high-risk and General Purpose AI Systems;
- II. Fault-based liability with a reversed burden of proof for other AI systems;
- III. Specifying immaterial damages and societal harms; and,
- IV. Increased clarity on AI liability across the value chain.

2 R. A. Posner and W. M. Landes, 'The Positive Economic Theory of Tort Law', *Georgia Law Review*, Vol. 15, 1980, p. 851.

3 See e.g. L. Kaplow and S. Shavell, *Economic Analysis of Law*, in A.J. Auerbach and M. Feldstein (eds.) *Handbook of Public Economics*, Vol.3, 2002, North Holland Publishing.

4 A. Galasso and H. Luo, *How does product liability risk affect innovation? Evidence from medical implants*, 2018, [http://individual.utoronto.ca/galasso/research\\_files/Papers/GalassoLuo\\_2July2018.pdf](http://individual.utoronto.ca/galasso/research_files/Papers/GalassoLuo_2July2018.pdf). See also, L. Kaplow, 'Rules Versus Standards: An Economic Analysis', *Duke Law Journal*, Vol. 42, 2012, pp. 557-629.

5 European Parliament Research Service, *Civil liability regime for artificial intelligence, European added value assessment*, September 2020, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654178/EPRS\\_STU\(2020\)654178\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654178/EPRS_STU(2020)654178_EN.pdf).

6 Proposal for a Directive of the European Parliament and of the Council on liability for defective products, COM(2022) 495 final, 28.9.2022.

7 See recital 3 of the proposed AILD.

## Recommendations

### I. Strict liability for high-risk and General Purpose AI Systems

FLI recommends a strict liability regime for high-risk AI systems as defined by the AI Act,<sup>8</sup> as well as for General Purpose AI Systems.<sup>9</sup> Strict liability has been the solution for comparable risk situations in the past and accounts for the knowledge gap between the operator<sup>10</sup> of a system and courts. In line with the thinking of the Expert Group on Liability and New Technologies and the European Parliament, strict liability appears to be the appropriate approach for dealing with risks stemming from high-risk and General Purpose AI Systems.

Fault-based liability applies in scenarios when a duty of care is breached. For cases related to AI, the necessary level of care and acceptable level of risk may be difficult to determine in view of its capabilities and rapidly evolving nature.<sup>11</sup> Both levels are ultimately determined by the judiciary, which is likely to have less technological and risk knowledge than developers and manufacturers, ultimately leading to a lack of legal certainty.<sup>12</sup>

Fault-based liability can be ineffective when risks are not well-understood based on the state of the scientific and technical knowledge. For instance, while a General Purpose AI System used as a chatbot in a healthcare context has great potential benefits in increasing access to relevant health information, it could equally become a source for potential damage by encouraging self-harm.<sup>13</sup> In such cases, courts may lack the scientific and technical knowledge to establish the appropriate level of care, or to judge whether the damage was caused due to a negligent action or omission.

As opposed to fault-based liability, a strict liability regime can be an instrument of technology risk control under uncertainty because there is no need for in-depth knowledge on the optimal level of care.<sup>14</sup> This incentivises further development of technologies to make them safer and may increase acceptance by the public.<sup>15</sup> In this sense, strict liability covers “actions that are fundamentally desired by society, and for which the appropriate incentives should be provided.”<sup>16</sup>

Indeed, risks emerging from new technologies (e.g. means of transport, energy or pipelines) have frequently been regulated by a strict liability regime in the past<sup>17</sup>. Such liability rules cover specific risks linked to objects or activities deemed permissible, but with a residual risk of harm.<sup>18</sup> In particular, situations that trigger strict liability can be divided into the following:<sup>19</sup>

- 
- 8 Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM/2021/206 final, 21.04.2021.
- 9 FLI understands a General Purpose AI System as a technology that can accomplish or be adapted to accomplish a range of distinct tasks, including some for which it was not intentionally and specifically trained. See Carlos I. Gutierrez, Anthony Aguirre, Risto Uuk, “The European Union could rethink its definition of General Purpose AI Systems (GPAIS)”, OECD AI Policy Observatory, 7 November 2022, <https://oecd.ai/en/wonk/eu-definition-gpais>.
- 10 The term “operator” should be understood in a broad sense. This would include “the provider, the user, the authorised representative, the importer and the distributor” as defined by the AI Act. However, it could also include additional natural or legal persons who exercise a degree of control over the risk in line with the European Parliament’s resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)).
- 11 Buiten, Miriam and de Streel, Alexandre and Peitz, Martin, EU Liability Rules for the Age of Artificial Intelligence (April 1, 2021). Available at SSRN: <https://ssrn.com/abstract=3817520> or <http://dx.doi.org/10.2139/ssrn.3817520>.
- 12 Zech, H. Liability for AI: public policy considerations. ERA Forum 22, 147–158 (2021). <https://doi.org/10.1007/s12027-020-00648-0>.
- 13 <https://www.nabla.com/blog/gpt-3/>.
- 14 Buiten, Miriam and de Streel, Alexandre and Peitz, Martin, EU Liability Rules for the Age of Artificial Intelligence (April 1, 2021). Available at SSRN: <https://ssrn.com/abstract=3817520> or <http://dx.doi.org/10.2139/ssrn.3817520>; Zech, H. Liability for AI: public policy considerations. ERA Forum 22, 147–158 (2021). <https://doi.org/10.1007/s12027-020-00648-0>.
- 15 Zech, H. Liability for AI: public policy considerations. ERA Forum 22, 147–158 (2021). <https://doi.org/10.1007/s12027-020-00648-0>.
- 16 Ibid.
- 17 European Commission, Directorate-General for Justice and Consumers, Liability for artificial intelligence and other emerging digital technologies, Publications Office, 2019, <https://data.europa.eu/doi/10.2838/573689>.
- 18 Ibid.
- 19 European Parliament Research Service, Civil liability regime for artificial intelligence, European added value assessment, September 2020, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654178/EPRS\\_STU\(2020\)654178\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654178/EPRS_STU(2020)654178_EN.pdf).

- Strict liability for damage caused by things (applicable to holders or owners of weapons, explosives, motor vehicles, or collapsing buildings);
- Strict liability for damage caused by dangerous activities (applicable to those engaging in hunting, diving, fireworks, tree felling, or use of nuclear power);
- Strict liability for damage caused by animals (applicable to owners or keepers of animals, which may or may not be limited to certain types of animals); and
- Vicarious liability (applicable in situations when a person is liable for an action of another person, such as liability of a principal for an agent, liability of a parent for a child, and liability of an employer for an employee).<sup>20</sup>

The Expert Group on Liability and New Technologies concluded that “emerging digital technologies that may typically cause significant harm comparable to the risks already subject to strict liability should also be subject to strict liability.”<sup>21</sup> A strict liability regime was also called for by the European Parliament in its own-initiative resolution from October 2020<sup>22</sup> and was highlighted as a preference in the Commission’s public consultation.<sup>23</sup> The Commission acknowledged these preferences in the explanatory memorandum accompanying the proposal,<sup>24</sup> yet it chose a different approach, thereby decreasing the level of consumer protection while increasing the burden on claimants.

In this context, a strict liability regime should equally apply to General Purpose AI Systems given their emerging capabilities, surprising and unexpected outputs, instrumental autonomous goal development and low level of interpretability<sup>25</sup>. Furthermore, their frequent use at the source of the value chain often leads to significant downstream risks if liability regimes do not create the right incentives for developers.

## II. Fault-based liability with a reversed burden of proof for other AI systems

FLI recommends that AI systems other than high-risk and General Purpose AI Systems fall under a fault-based liability regime where the presumption of fault lies on the operator. Pursuing this course of action would ease the burden for claimants and facilitate their access to justice by minimising their information asymmetry and transaction costs. Operators can rebut this presumption of fault by proving their observance to the required level of care, and that the damage caused was not their fault.

Fault-based liability is liability for damage caused intentionally or by a negligent act or omission. Fault-based liability is triggered when a duty of care is breached. A fault-based liability regime should apply to AI systems not covered by strict liability, i.e. systems that do not fall under the category of high-risk or General Purpose AI systems. Fault-based liability normally requires the claimant to prove the existence of a damage, a fault (breach of a duty of care) by the defendant as well as the causality link between the damage and the fault.

20 It should be noted that liability rules currently diverge across the EU and that the above-mentioned scenarios are merely examples of strict liability regimes.

21 European Commission, Directorate-General for Justice and Consumers, Liability for artificial intelligence and other emerging digital technologies, Publications Office, 2019, <https://data.europa.eu/doi/10.2838/573689>.

22 European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)), [https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html).

23 [https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12979-Civil-liability-adapting-liability-rules-to-the-digital-age-and-artificial-intelligence/public-consultation\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12979-Civil-liability-adapting-liability-rules-to-the-digital-age-and-artificial-intelligence/public-consultation_en).

24 “Various national legal systems provide for different strict liability regimes. Elements for such a regime at Union level were also suggested by the European Parliament in its own-initiative resolution of 20 October 2020, consisting of a limited strict liability regime for certain AI-enabled technologies and a facilitated burden of proof under fault-based liability rules. The public consultations also highlighted a preference for such a regime among respondents (except for non-SMEs businesses), whether or not coupled with mandatory insurance. However, the proposal takes into account the differences between national legal traditions and the fact that the kind of products and services equipped with AI systems that could affect the public at large and put at risk important legal rights, such as the right to life, health and property, and therefore could be subject to a strict liability regime, are not yet widely available on the market.”

25 J. Wei et al., “Emergent abilities of large language models,” arXiv preprint arXiv:2206.07682, 2022. and D. Ganguli et al., “Predictability and surprise in large generative models,” 2022, pp. 1747–1764.

As acknowledged by the Commission, several specific characteristics of AI - such as their autonomous behaviour, continuous adaptation, limited predictability, and opacity - can make it difficult and costly for injured parties to identify and prove the fault of a potentially liable entity, and receive compensation.<sup>26</sup> Specifically, harmed individuals are subject to significant information asymmetry with respect to the AI systems they interact with because they may not know which code or input caused a harm. The interplay between different systems and components, the multitude of actors involved as well as the increasing autonomy of AI systems adds to the complexity in proving fault.<sup>27</sup> This is why a number of experts,<sup>28</sup> as well as the European Parliament,<sup>29</sup> have called for a rebuttable presumption of fault of the operator, in addition to a rebuttable presumption of a causal link between the damage and the fault.

FLI believes that a fault-based liability regime with a reversed burden of proof for AI systems other than high-risk and General Purpose AI Systems is a reasonable and balanced approach. Following the risk-based approach of the AI Act, it seems sensible to have less stringent requirements for these AI systems as compared to high-risk and General Purpose AI Systems, which should be subject to a strict liability regime. This is also in line with the tradition to reserve strict liability for particularly dangerous activities or things.

### III. Specifying immaterial damages and societal harms

FLI calls for a harmonisation of compensable damages across the EU that should include immaterial and societal harms. Covering immaterial harms is necessary to account for the particular nature of damages caused by AI systems. Furthermore, to ensure that victims harmed by AI systems are equally protected no matter which EU country they live in, the specific types of recoverable damages should be defined by EU law.

The importance of immaterial harms caused by AI systems was recognised in the Commission's 2020 White Paper on Artificial Intelligence.<sup>30</sup> It specifically lists the "loss of privacy, limitations to the right of freedom of expression, human dignity, discrimination for instance in access to employment" amongst the harms. The proposed AILD allows for immaterial damages to be covered, such as discrimination, privacy infringements, other breaches of fundamental rights or cybersecurity vulnerabilities, which is a welcome extension and reflects the distinct nature of harms stemming from AI systems.

However, the proposed Directive leaves it up to Member States to define through their national laws the exact types of damages that will be covered.<sup>31</sup> As opposed to the AILD, the Commission's PLD proposal does include a harmonised definition of what constitutes damage<sup>32</sup>. It follows that people harmed by defective products will enjoy uniform protection across the EU, while the protection of people harmed by AI systems will depend entirely on Member States' legislation.

This could mean that a person discriminated against by a credit scoring AI system could claim damages for such discrimination in one Member State but not in another. It is also not clear whether

26 European Commission (2021), Civil liability – adapting liability rules to the digital age and artificial intelligence, Inception Impact Assessment, [https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12979-Civil-liability-adapting-liability-rules-to-the-digital-age-and-artificial-intelligence\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12979-Civil-liability-adapting-liability-rules-to-the-digital-age-and-artificial-intelligence_en).

27 Buiten, Miriam and de Streef, Alexandre and Peitz, Martin, EU Liability Rules for the Age of Artificial Intelligence (April 1, 2021). Available at SSRN: <https://ssrn.com/abstract=3817520> or <http://dx.doi.org/10.2139/ssrn.3817520>; Zech, H. Liability for AI: public policy considerations. ERA Forum 22, 147–158 (2021). <https://doi.org/10.1007/s12027-020-00648-0>.

28 European Commission, Directorate-General for Justice and Consumers, Liability for artificial intelligence and other emerging digital technologies, Publications Office, 2019, <https://data.europa.eu/doi/10.2838/573689>; Buiten, Miriam and de Streef, Alexandre and Peitz, Martin, EU Liability Rules for the Age of Artificial Intelligence (April 1, 2021). Available at SSRN: <https://ssrn.com/abstract=3817520> or <http://dx.doi.org/10.2139/ssrn.3817520>; Zech, H. Liability for AI: public policy considerations. ERA Forum 22, 147–158 (2021). <https://doi.org/10.1007/s12027-020-00648-0>.

29 European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)), [https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html).

30 European Commission, White Paper On Artificial Intelligence - A European approach to excellence and trust, COM(2020) 65 final, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf).

31 See Article 2(9) and recital 22 of the proposed AILD and its Explanatory Memorandum.

32 See Article 4(6) of the proposed PLD.

the Directive would allow for broader societal harms caused by AI systems to be covered, such as manipulation at scale, election interference or environmental harms.

Indeed, it appears there is currently no uniform approach across the EU when it comes to the recoverability of non-material damages. One category of EU countries does not differentiate between material and immaterial damage and considers both damages equally recoverable. This category includes Belgium, France, Luxembourg, Spain, Hungary and Slovenia. The second category departs from the premise that immaterial damage is generally non-recoverable and only allows recovery of immaterial damage if expressly provided for by law. This approach is followed by countries like Italy, Germany, Poland, Austria, the Netherlands, Estonia, Lithuania and the Nordic Countries.<sup>33</sup>

#### IV. Increased clarity on AI liability across the value chain

FLI calls for clear legal rules in the AILD to allocate liability for damages caused by AI systems across the value chain. It is crucial for persons harmed by AI systems to know who can be held accountable, and given the complexity of AI systems including their value chains, the AILD must clarify which operators can be held liable.

Numerous parties are involved in the design, development, and deployment of AI systems (e.g. hardware manufacturers, software developers, data suppliers, providers of network services, deployers and employees). From the perspective of a victim harmed by an AI system, it is indispensable to know who can be held accountable. Yet, identifying the liable party in the face of the complexity of this value chain is a challenge and in some cases it will be impossible for a victim to identify the responsible actor.<sup>34</sup>

Nevertheless, the proposed AILD remains silent with respect to the allocation of liability for damages caused by AI systems. The proposal merely uses the term “defendant” as well as the concepts of “provider” and “user” as defined by the AI Act, but it does not specify which of these operators can be held liable and avoids allocating liability to specific actors across the value chain.

In contrast to the proposed AILD, Article 7 of the proposed PLD does establish rules on which economic operators are to be held liable for defective products. In this context, the primary liable party is the manufacturer, the manufacturer of a component or the importer. Additional parties can be held liable under Article 7, such as distributors or any natural or legal person that modifies a product.

In the same vein, a harmonised approach at the EU level with clear legal rules allocating liability also needs to be established for the AILD. As called for by the Expert Group on Liability and New Technologies,<sup>35</sup> the person held liable should be the operator, meaning the person in control of the risk and who benefits from the operation. In the case of more than one operator, e.g. a frontend and a backend operator, the operator with the higher level of control over the risks should be the one held liable with the possibility for redress claims between them. A system of joint and several liability could also be considered.<sup>36</sup> Further inspiration can be taken from the European Parliament’s 2020 Resolution,<sup>37</sup> which made an attempt to specify rules on the apportionment of liability, coupled with rules for joint liability, recourse for compensation, and liability insurance.

33 Commission Staff Working Document, Impact Assessment Report accompanying the document Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence, SWD(2022) 319 final, 28.9.2022, [https://ec.europa.eu/info/sites/default/files/1\\_4\\_197608\\_impact\\_asse\\_dir\\_ai\\_en.pdf](https://ec.europa.eu/info/sites/default/files/1_4_197608_impact_asse_dir_ai_en.pdf).

34 See recital 60 of the AI Act; See also Benhamou, Yaniv & Ferland, Justine. (2020). ARTIFICIAL INTELLIGENCE & DAMAGES: ASSESSING LIABILITY AND CALCULATING THE DAMAGES.

35 European Commission, Directorate-General for Justice and Consumers, Liability for artificial intelligence and other emerging digital technologies, Publications Office, 2019, <https://data.europa.eu/doi/10.2838/573689>.

36 Buiten, Miriam and de Streef, Alexandre and Peitz, Martin, EU Liability Rules for the Age of Artificial Intelligence (April 1, 2021). Available at SSRN: <https://ssrn.com/abstract=3817520>.

37 European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)), [https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html).