Cambridge Forum on AI: Law and Governance (2025), **1**, e5, 1–23
doi:10.1017/cfl.2024.8

RESEARCH ARTICLE

Mapping Generative AI rules and liability scenarios in the AI Act, and in the proposed EU liability rules for AI liability

Teresa Rodríguez de Las Heras Ballell

Private Law Department, Universidad Carlos III de Madrid, Getafe, Spain
Email: teresa.rodriguezdelasheras@uc3m.es

Paper prepared in the framework of the Project funded by the BBVA Foundation in the 2021 Call for BBVA Foundation Grants for Scientific Research Projects, in the area of Social Sciences, of which the author is PI, entitled *Responsible Algorithms. Development of a European regulatory framework for the responsible automation of decision-making and contractual relations/Algoritmos Responsables. Desarrollo de un marco normativo europeo para la automatización responsable de la toma de decisiones y las relaciones contractuales* (2022–2024).

(Received 2 August 2024; revised 1 November 2024; accepted 8 November 2024)

Abstract

The Paper aims to explore whether the current ecosystem of existing and still-to-be adopted rules on artificial intelligence (AI) systems in the European Union does fully and adequately address the liability for damages caused by Generative AI system. It maps first and primarily the distinctive features and functional characteristics of Generative AI likely to impact on regulatory and legal considerations and, in particular, on determining the specific regulatory regime governing Generative AI and on tracing and allocating liability along the value chain pursuant to the AI Act. On the basis of this mapping exercise, the Paper focuses on testing the liability rules as provided for the draft Artificial Intelligence Liability Directive and the Revised Product Liability Directive and assessing their sufficiency and effectiveness in the face of Generative AI. Beyond the assessment of the rules laid down in the above-referred texts, the Paper briefly describes other liability scenarios to be explored in future works.

Keywords: Generative AI; general-purpose AI model; AI system; product liability; AI Act; large language models

1. Setting the scene: Policy and regulation of AI in the EU

Artificial intelligence (AI) has burst onto the global scene, arousing equal parts disquiet and fascination. The exponential development of AI and, in particular, its irruption in the market and the expansion of its uncountable uses with astonishing decision-making capabilities and an impressive content-generation potential have been accompanied by a growing and, sometimes alarming, perception of risks – performance risks, security risks, control risks, ethical risks, economic risks, societal risks (World Economic Forum, 2018) – that has marked the regulatory debate, the public opinion and the social acceptance (European Commission, 2021) of its advances. Hence, the attention of legislators and regulators on AI has spotlighted the critical (and pressing) need to reconcile its promises and its perils.

In such a context, the European Union (EU) has taken a leading role in its attempts to formulate and establish an ambitious legal framework for AI. Although the AI Act[1] is clearly the flagship regulation, it is neither the only one nor a comprehensive legal regime for AI. The AI Act is a key component of the broader and more complex ecosystem of (existing and prospective) pieces of legislation aimed at govern the development, deployment and use of AI systems in or in connection with the EU. In particular, the AI Act does not provide for a body of liability rules for damages caused by AI systems.

Alongside the regulatory response that has crystallized in the adoption of the AI Act, aimed to prevent and mitigate potential risks, the EU has paid particular attention to the adequacy and sufficiency of liability systems to confront AI damages. For various reasons, an AI system may be involved in causing damages (defects, wrong design, inaccurate or biased data, poor training, evitable or inevitable hacking attacks, human mistakes, unexpected learning process) to individuals or property. Where the risks materialize, the challenge is to determine whether existing liability rules are adequate, effective and sufficient to compensate in an equivalent way (compared to a case where damages are caused without the use of AI) the victims. As a matter of fact, promoting safe, reliable and high-quality AI in Europe had become one of the backbones of the EU Digital Strategy defined in the strategic package adopted on 19 February 2020. The White Paper *Artificial Intelligence – A European approach to excellence and trust*[2] and the Report *on the safety and liability implications of AI, IoT and Robotics*[3] defined the coordinates for the Europe's digital future (*Shaping Europe's Digital Future,* European Commission, 2020). Since the formation in 2018 of the *Expert Group on Liability and New Technologies* divided into two formations[4]: New Technologies formation and Product Liability formation, there has been a succession of reports, working papers and legislative initiatives of various nature to explore precisely the AI terrain with the existing liability tools and test their adequacy and effectiveness.

Mindful of the distinctive features of AI systems (complexity, opacity, vulnerability, openness, data dependence, autonomy) and their potential impact on the effective application of pre-existing liability regimes, the EU has then embarked on a review of the existing liability rules to assess their adequacy and effectiveness in the face of the intensive and extensive use of AI. Such an assessment resulted into the publication of two proposals of Directive on the 28th of September of 2022: the Proposal for a Directive on the adaptation of non-contractual fault-based liability rules to artificial intelligence[5] (AILD) and the Proposal for a revision of the Directive of the European Parliament and of the Council on liability for defective products[6] (revPLD) that are key pieces in the legislative review effort to shape a comprehensive and coherent liability regime for damages caused by AI systems (Rodríguez de Las Heras Ballell, 2023). While the AILD is still an in-progress proposal, the revPLD was approved by the European Parliament on the 12th of March 2024.[7]

[1]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized 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), PE/24/2024/REV/1, *OJ L, 2024/1689, 12.7.2024.*

[2]COM(2020) 65 final, Brussels, 19.2.2020.

[3]COM(2020) 64 final, Brussels, 19.2.2020.

[4]Published in the register of expert groups on 9 March 2018 – https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&do=groupDetail.groupDetail&groupID=3592. The expert group's mission was to

> provide the Commission with expertise on the applicability of the Product Liability Directive to traditional products, new technologies and new societal challenges (Product Liability Directive formation) and assist the Commission in developing principles that can serve as guidelines for possible adaptations of applicable laws at EU and national level relating to new technologies (New Technologies formation).

[5]COM/2022/496 final.

[6]COM/2022/495 final.

[7]COM(2022)0495 – C9-0322/2022 – 2022/0302(COD).

These texts interplay with the AI Act as well as with the national rules on fault-based liability in setting out a liability system for AI system that, concurrently, needs to be complemented with a constellation of existing rules in the fields of contract law, privacy, intellectual property rights or competition law, *inter alia*.

In this vast and multifaceted policymaking and lawmaking process, the advent and the momentum gained by Generative AI irrupted into the regulatory scene signaling a significant paradigm shift in the AI landscape. Throughout the deliberative process of the AI Act the attention paid to Generative AI and its implications crystallized in various policy decisions and drafting solutions. The notion of AI system has been carefully refined and new definitions were added to the draft AI Act to ensure that it was suited to address the challenges of these emerging models. From the initial inclusion of the term of "foundation models" in intermediate versions of the text to the final incorporation of the concept of "general-purpose AI model" (and system), the AI Act has struggle to embrace the extraordinary potential of Generative AI.

In addition to these terminological and conceptual efforts to accommodate the scope of application to Generative AI, the AI Act does also provide for specific solutions designed for these models.

Nonetheless, there are legal and regulatory implications of Generative AI that may exceed the proposed solutions. Understanding the extent of the paradigm shift that Generative AI means in the AI landscape will allow to assess the sufficiency of the measures adopted and to identify possible shortcomings and gaps in the current EU framework.

In particular, Generative AI raises specific problems in the compliance of the AI Act obligations and in the application of liability rules that have to be acknowledged and properly addressed. The multimodality, the emergence factor, the scalability or the generality of tasks may mismatch the assumption underlying the obligations and requirements laid down for AI systems.

Besides, the generative capabilities of these systems producing content of any kind, text, images, audiovisual or functional information (derivative AI systems) exacerbate certain liability risks inherent to AI systems or tend to trigger specific damages related to the infringement of privacy or IP rights, provoked by the reliance on the outputs while assisting users in a diversity of decision makings, or caused by the harming effects of the outputs themselves on affected persons.

The Paper aims to explore whether the current ecosystem of existing and still-to-be adopted rules on AI systems does fully and adequately address the distinctive features of Generative AI, with special consideration to the interaction between the AI Act and the liability rules as provided for the draft AILD and the revPLD. Beyond the assessment of the rules laid down in the above-referred (draft) Directives, the Paper briefly lists other liability scenarios to be explored in future works. With such an aim, the Paper is structured as follows. Part 2 traces how Generative AI has been included within the scope of application of the AI Act and which specific rules have been laid down accordingly. Part 3 focused on the liability rules as recently modified in the EU with special consideration to AI: the proposal of AILD and the revised PLD. Part 4 maps other potential liability scenarios and concludes with final remarks.

2. Generative AI in the AI Act: Definition, relevant features and rules

The notion of AI system is the mainstay of the EU regulatory strategy on AI. Accurately defining the concept and, more importantly, properly identifying its distinctive and differential features underlying the rationale for a specific regulation are instrumental to deploy the regulatory strategy in an effective and justified way. The definition of AI systems does not only demarcate the scope of application of the regulatory framework, but foremost it must capture its distinctive characteristics and encapsulate the risk triggers that require regulatory action.

The legal definition of AI system is neither a simple task nor an innocuous exercise. Not surprisingly, the definition has evolved significantly in the development of the AI Act from its initial

wording as laid down in the proposal. Along such an evolution, the momentum of Generative AI has impacted on the conceptualization process and convulsed the terminological options and the drafting decisions in the AI Act. In tracing the drafting evolution throughout the successive versions of the text, meaningful findings are revealed.

2.1 Tracing the evolution of the definition of AI systems in the AI Act: Meaning and significance

The formulation of a definition of AI systems for legal purposes faces a number of challenges. It must be free from technological determinants and maintain sufficient (technological) neutrality to encompass the various solutions available (or to come) on the market and to avoid obsolescence by being able to integrate future technological developments. It must be translated into functional and regulatory-relevant features, operational characteristics and technical specifications. The evolution of the definition of an AI system in the European proposal for an AI Act[8] highlights these difficulties.

The first definition proposed in the AI Act[9] immediately raised several questions and aroused some criticism. First, it failed to convincingly and decisively delineate the proposed AI systems to be regulated in contrast with the already commonly used and widely known computer programs (software). The regulatory requirements imposed by the new text required a clear and objective distinction as to their scope of application. Second, the reference to certain techniques listed in an Annex, despite the establishment of a review mechanism, questioned (in addition to the uniqueness of the chosen techniques themselves)[10] the technological neutrality, the adaptive capacity to new solutions and the very soundness of a purely descriptive definition incapable of offering a functional concept.

The proposed definition evolved[11] aligning itself with the OECD notion (OECD, 2019a, 2019b) which, following a subsequent revision (OECD, 2023a),[12] strengthens some of the differential functional features and qualifies others (OECD, 2024) in order to accommodate the most recent developments that had burst onto the international scene and media debate (large language models, Generative AI, general AI). Enlarging the definition of AI systems so as to include emerging paradigms was the first drafting option. Nonetheless, it was not implemented on an isolated manner, but in conjunction with the addition of new definitions aimed at embracing the emerging models. Thus, for the purposes of the AI Act, in order to ensure its coverage in the scope of application and to be able to differentiate, in turn, different regulatory regimes depending on the AI model, in the

[8] COM/2021/206 final.

[9] The proposed AI Regulation (Art.3(1)) defined the AI system as:

> software that is developed using one or more of the techniques and strategies listed in Annex I and that can, for a given set of human-defined objectives, generate output information such as content, predictions, recommendations or decisions that influence the environments with which it interacts.

[10] The proposed AI Regulation, Annex I, Artificial Intelligence Techniques referred to in Article 3, point 1:

> Machine learning strategies, including supervised, unsupervised and reinforcement learning, employing a wide variety of methods, including deep learning. Logic and knowledge-based strategies, especially knowledge representation, inductive (logic) programming, knowledge bases, inference and deduction engines, expert and (symbolic) reasoning systems. Statistical strategies, Bayesian estimation, search methods and optimisation.

[11] In the version of the text dated 14 July 2023. (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))1 defines an "AI system" as *a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations or decisions that influence physical or virtual environments*. The subsequent version of the text, after the compromise agreement reached was made available on 24 January 2024 and includes the following definition: *An AI system is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments.*

[12] *An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input 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 adaptiveness after deployment*.

successive amendments to the proposed AI Act two other definitions were added along with the notion of AI system that were not in the initial version: "foundation model"[13] and "general purpose AI system."[14] Subsequently, in the subsequent texts, these definitions have also changed.

Yet, in the approved version and after the final corrigendum,[15] the Regulation pivots on a notion of AI systems[16] aligned with the most recently updated definition recommended by OECD (OECD, 2023a, 2023b, 2024) and two other concepts: general-purpose AI model[17] and general-purpose AI system.[18]

By following the definition of AI system proposed by OECD, the AI Act ensures coordination with international standards and converges with the global consensus that the recommendation reflects.

The concept of AI system is articulated along four main axes: interactivity, adaptiveness, autonomy and influence on the environment. The potential of AI systems to (relatively but increasingly) autonomously learn (Anderljung et al, 2023; Ngo et al, 2023) according to varying levels of autonomy challenges settled anthropocentric notions of consent and intent, error, fault or negligence. The very rapid advances of generative AI models and the exponential growth of their capabilities are already beginning to raise the possibility that they may derive in "emergent behaviors" (Chan et al., 2023; Perez et al., 2023) aimed at circumventing human control, optimizing resources to achieve the goal in a sub-optimal way, using persuasion techniques or pretending to be human. Adaptive, learning and evolving capacity injects unpredictability into the outcome and raises questions concerning the

[13] In the text published on 24 of January 2024, the definition of "general purpose AI model" is:

> an AI model, including when trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications. This does not cover AI models that are used before release on the market for research, development and prototyping activities.

Previously, in the P9_TA(2023)0236 (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))1 Art.3.1.c)

> "foundation model" means an AI model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks.

The definition is not included in the latest text after the political agreement in December 2023 and as published in January 2024.

[14] P9_TA(2023)0236 (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))1 Art. 3.1.d).

> "General purpose AI system" means an AI system that can be used in and adapted to a wide range of applications for which it was not intentionally and specifically designed.

[15] P9_TA(2024)0138. (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)). 19.4.2024.

[16] Art. 3(1) AI Act:

> "AI system" means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments.

[17] Art. 3(63) AI Act:

> "general-purpose AI model" means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market.

[18] Art. 3(66) AI Act:

> "general-purpose AI system" means an AI system which is based on a general-purpose AI model and which has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems.

treatment of error, the effect of unexpected or surprising learning, the attribution of legal effects or the allocation of liability. Yet, AI systems are capable of generating outcomes that influence the environment in which they operate on the basis of a set of objectives, either explicit or implicit, which may have been determined at design (Raji et al., 2022) or learned later in their operation. These outcomes may consist not only of predictions or recommendations but also of decisions and content of the most diverse nature. This is widely considered in generative AI models. To generate these results, systems use data or information that they receive, infer, perceive from the environment and learn by means of learning methods.

Therefore, the definition of AI system in the AI Act does, on the one hand, embody the disruptive features on which the legislative action is founded, as well as, on the other hand, it shows certain receptiveness to the specific and additional challenges posed by Generative AI. But, in addition to this permeability of the "AI system" concept to various functional and distinctive characteristics of Generative AI, the AI Act adds the definition of general-purpose AI models (and general-purpose AI system) to properly embrace the paradigm shift that large generative AI models epitomize.[19]

The introduction of these definitions is not a simple terminological wink toward the emerging models, but it is accompanied by other legal and regulatory implications – in the personal scope of application of the AI Act, in the provision for specific obligations, in the application of rules for models posing systemic risks. Nevertheless, the accommodation of the AI Act's and other correlated rules' logic to the functional characteristics of Generative AI models may not be perfect and complete. Certain gaps and imperfections may be spotted.

2.2 "Decoding" Generative AI models in the context of the AI Act: Relevant features

Generative AI models are covered in the AI Act under the notion of general-purpose AI models. After the terminological swinging between possible definitory options (from "foundation models" to "general-purpose AI models or systems") successively explored in the chain of versions of the text, the AI Act does not define "generative AI models" but rather utilize them as typical example of "general-purpose AI models" that is the term finally adopted as a defined concept in the Regulation.

In the choice among the alternative terms, different features are purportedly emphasized. A term is preferred and chosen so as to stress either the basal character of these models as essential components of multiple AI systems (foundation models) – the word "foundation" specifies the role these models play: a foundation model is itself incomplete but serves as the common basis from which many task-specific models are built via adaptation¡ (Bommasani et al., 2021), or the multimodality (Large Multimodal Models), or the variety of possible uses (general-purpose AI models), or even the severity and seriousness of the potential risks to society, global security and public safety – models posing systemic risks or frontier models (Bommasani et al., 2021; Nerlich, 2023). Beyond the preferences for one term or another, all of them do partially or totally reflect the distinctive features and the functional characteristics of Generative AI. In understanding the paradigm shift that this class of models embodies, and assessing the regulatory and legal implications, all these functional characteristics become relevant (Hacker et al., 2023).

One of the most visibly differential features of Generative AI, as the chosen terminology bluntly unveiled (general-purpose AI models), is generality. Advanced models can perform a broad array of tasks. Unlike task-specific, or purpose-specific AI system, task-agnostic models, as they are also described, can serve multiple downstream tasks. As a matter of fact, the possibility that a single model

[19] Recital 97 AI Act:

> The notion of general-purpose AI models should be clearly defined and set apart from the notion of AI systems to enable legal certainty. The definition should be based on the key functional characteristics of a general-purpose AI model, in particular the generality and the capability to competently perform a wide range of distinct tasks.

can perform and serve for a wide range of tasks is deemed to mark the beginning of the era of foundation models (Bommasani et al., 2023). The generality and the versatility of these models are perfectly staged by the performance of advanced language models that prove to be adapted to any downstream task in response to a provided prompt, where the prompt represents a description of the task in natural language. Impressively, the prompt triggers an emergent property that the model has not trained for and, even in some cases, it can be affirmed that the task was not anticipated to arise.

The feature of generality and diversity of performed tasks has substantial implications in the regulatory logic. A model that can perform a variety of tasks, even unanticipated ones, disrupts the risk-based approach underlying the regulatory strategy that the AI Act embodies. The risk classification is linked to the purpose, the use or the application. Both in the classification of prohibited AI practices (Art. 5 AI Act) as well in the list of high-risk AI system in Annex III as referred to in Article 6(2) AI Act, the use or the intended use is pivotal to the application of the regulatory framework. General-purpose AI models provoke the fracture of this risk-based purpose specific regulatory model. That is an obvious challenge. Therefore, the integration of multiple-purpose AI models in the AI Act was a crucial step forward that required important efforts to ensure coherence and consistency.

Assuming unanticipated outputs and/or unexpected or unlikely effects arising from tasks or uses the model has not specifically trained for has strong implications in the liability narrative too. Causality is weakened, while the foreseeability of consequences may be justifiably deemed rather remote.

The definition of general-purpose AI model and of general-purpose AI system in the AI Act (Art. 4(63) and (66) AI Act) acknowledges the feature of generality – "displays significant generality and is capable of competently performing a wide range of distinct tasks" and "has the capability to serve a variety of purposes." Should the purpose be general, multiple or unanticipated, the risk classification does not work. Therefore, the AI Act needs to find a proper and coherent approach to embrace and address the risks of these models.

Compliance requirements based on intended purposes may fail or prove to be unsuited to Generative AI. Random outputs and unintended actions obscure the efforts to predict and anticipate concrete impacts. Hence, risk management systems and transparency obligations may lose effectiveness in the landscape of Generative AI (Novelli et al, 2024b).

The idea of multiple possible tasks to be performed by the model places Generative AI in a challenging position for liability purposes. It instils a disconcerting level of unpredictability not only as for the harmful potential of a specific output but on the entire extent, risk involved and type of tasks that the model turns out to perform.

From the perspective of the AI supply chain, general-purpose AI models are essential components of AI systems. As the descriptive term of "foundation" reveals, they play a key role by serving as the common basis from which many task-specific models are built via adaptation. Thus, these models are further fine-tuned and adapted to multiple applications. The power of "foundations models" is precisely their potential to grow, expand and be integrated into a number of AI systems. That powers a strong homogenization capability. Concurrently, scalability and scale are also triggers of possible viral risks. In operating as the bedrock of subsequent applications and further development of AI systems throughout the supply chain, defects, biases or vulnerabilities of few foundation models might be inherited by and span by all AI systems based on them.

Such a scale effect of general-purpose AI models is acknowledged by the AI Act with the categorization of models that can pose systemic risks. The notion of systemic risk is underpinned by the ideas of criticality, adverse effects, extent of risks, replicability potential, public safety and global security. Thus, systemic risks comprise any actual or reasonably foreseeable negative effects in relation to, among others, disruptions of critical sectors; serious consequences to public health and safety; threats to democratic values, public and economic security; serious biases that may harm individuals or communities; severe privacy infringements or promotion of misinformation or disinformation (Recital 110 AI Act).

In a liability-centered reading, foundations models can become the point of failure in the AI supply chain, as a failure in the foundation model compromises the entire value chain. Concurrently, they may act as sources of virality while replicating a defect in downstream application or "infecting" AI systems integrating the model with biases or harmful failures.

On a technical level, general-purpose AI models are based on the idea of transfer learning (Bozinovski, 2020; Pratt et al., 1991). Under this training trend, the "knowledge" learned from one task (object recognition in images) is transferred and applied to another task (activity recognition in videos) and, as a result, the model learns to deploy its capabilities in a wide range of tasks, even in unexpected ones. To that end, scale is critical. Foundation models are trained on large amounts of data through various methods from supervised to unsupervised or reinforcement learning, but generally these models use self-supervision at scale. Subsequently, they are adapted and fine-tuned to a wide range of downstream tasks. These technical considerations are indeed pertinent and relevant for a legal analysis. Not surprisingly, a training model based on vast amounts of data raises challenges in the field of privacy and data protection. Whereas it also arouses concerns and disquiet as far as protection of copyright and other intellectual property rights are concerned.

Yet, as the term Generative AI prioritizes, these AI models show an extraordinary and impressive generative potential. The capability to generate content is moreover amplified by an expanding multimodality of these models. Many models (also defined as Large Multimodal Models) can process text, audio, image or video and generate outputs in various types of formats.

From a legal perspective, the generative power of these models is absolutely fascinating and challenging. This potential multiplies the sources of risks and diversifies the scenarios of liability, as the legal categorization of possible generated outputs may lead to rather diverse situations – an original creation, a scientific contribution, an offer to deal, a "deep fake," hatred speech, a settlement agreement, news, a misleading advertisement, a recommendation, etc. Therefore, in the attempt of mapping possible risk scenarios, and relevant liability regimes, the catalogue is ample and to a certain extent comprising of any plausible harmful effect on rights or interests. Not only multimodality provides high levels of versatility to the impact of the models' performance on the environment but also the most advance capabilities to imitate and model human language, creations and reactions exacerbate the vulnerability and the exposure of humans interacting with these models.

Furthermore, the significance of these models for the legal analysis is very much influenced by two elements. On the one hand, the notion of emergence. Emergent behavior instils substantial uncertainty about unanticipated outputs, but they do also raise fundamental legal questions ranging from the applicability of human-centric notions of creator or inventor to the attribution of liability (Boden, 2009; Grimmelmann, 2015, Sartor et al., 2018; Volokh et al., 2023). On the other hand, the risk of *hallucinations* (Ye et al., 2023) that challenges the effectiveness of the policy decisions and the regulatory requirements aimed to prevent and mitigate risks. Suited and effective strategies and methods to prevent harmful outputs need to be devised and implemented accordingly (Annex IX, Section 2, 2 on read teaming). Legal rules and regulatory requirements should be sensitive to these peculiarities and consistent with the methods that prove to be effective in generative AI models.

2.3 Rules on Generative AI in the AI Act: Unveiling the implications and assessing the effectiveness

General-purpose AI models fit uncomfortably into a regulatory framework based on a purpose-related risk approach. The AI Act struggled to render its incorporation coherent and less disturbing. The expansion strategy for the AI Act to embrace Generative AI has pivoted on three solutions. First, a limit expansion of the personal scope with special regard to the role/s played thereby in the AI supply chain. Second, a tiered risk classification parallel to and distinct from the purpose-specific risk classification underlying the AI Act. Third, a policy option for transparency-related requirements.

The connection between the general-purpose logic and the purpose-specific risk approach is not, however, ignored. Where the multiple-purpose potential crystallizes in a high-risk application the relevant rules are triggered and correspondingly apply throughout the AI supply chain. Thus, Article 25 AI Act provides for the responsibilities and obligations along the AI value chain to any distributor, importer, deployer or other third-party, who shall be considered to be a provider of a high-risk AI system for the purposes of the Act, among other circumstances, if they modify the intended purpose of an AI system, including a general-purpose AI system, which has not been classified as high-risk and has already been placed on the market or put into service in such a way that the AI system concerned becomes a high-risk AI system in accordance with Article 6 AI Act. So, the generality that defines general-purpose AI systems becomes compatible with a subsequent adaptation to the general purpose to a specific use classified as a high-risk one. Then, the risk-based machinery operating the AI Act can be put into motion allocating the regulatory obligations at the relevant "link" of the chain. Therefore, contextualizing the AI Act obligations along the AI supply chain is crucial.

In consistency with the foundation character of general-purpose AI models, their generality and the task-agnostic nature, the personal scope of the AI Act is proportionately and logically adapted. Unlike AI systems, only providers of general-purpose AI models are covered by the AI Act, while all other operators are linked to and qualified by activities relating to AI systems. This is consistent with the purpose of the Regulation as set out in Article 1 AI Act. In the second paragraph of the provision, point (e), it is specified that the Act establishes "harmonized rules for the placing on the market of general-purpose AI models." Since general-purpose AI models are essential components of AI systems, but do not in themselves constitute an AI system, the decisive determinant for the application of the AI Act is the placing on the market. This explains why the AI Act applies to providers of general-purpose AI models when they introduce them into the Union market.

The AI Act assimilates so the criticism (Hacker, 2023a, 2023b, 2021; Hacker et al., 2023) to previous versions of the text and solves the regulatory quandary between focusing exclusively on providers of general-purpose AI models or on other operators (as defined by Art. 3(8) AI Act). While the former policy decision leads to excessive and inefficient compliance obligations, and possibly to asphyxiating the AI supply chain, the latter one risks rendering compliance burdensome and equally ineffective as they depend upon information provided by the provider or lack resources and insights.

Providers of general-purpose AI models "head" the AI supply chain as their models may be the essential basis of a wide range of downstream systems and be fine-tuned for a number of varying downstream tasks. Thus, they play a singular role in the value chain that the AI Act translates into transparency measures – technical documentation and information as per the Annexes. Where these models are released under a free and open-source license (Art. 2(12) AI Act), the presumption that certain information is publicly available, and the transparency-related requirements are alleviated accordingly. Nonetheless, such exception is not absolute and complete. On the one hand, compliance is not excluded if the general-purpose AI model is considered to present a systemic risk (Art. 51 AI Act). Other the other hand, insofar as the assumed transparency inherent to a free and open-source licensing model does neither necessarily entail that substantial information on the data set used for the training or fine-tuning is revealed nor how the copyright law compliance is ensured, such obligations are to be complied with by such models as well.

This approach reveals that the AI Act reconstructs the risk classification for general-purpose AI models under a three-tier system that distinguishes between standard models, models released under free and open-source license and models posing systemic risks. Pursuant to this taxonomy, the AI Act

sets out relevant definitions (high-impact capabilities[20] and systemic risks[21]), additional obligations for systemic risk general-purpose AI models (Art. 55, in addition to the obligations provided for by Arts. 53 and 54) and a procedure to classify a general-purpose AI model with systemic risks, on the basis of its evaluated high-impact capabilities – with a presumption determined by the cumulative amount of computation used for its training measured in floating point operations and subject to thresholds, benchmarks and indicators that can be amended by delegated acts – or a decision of the Commission ex officio or following a qualified alert (Arts. 51 and 52 AI Act, and Annex XIII).

The primary transparency-based policy strategy deploys in obligations to provide technical documentation and information by general-purpose AI system providers downstream (Art. 53 and Annexes XI and XII) as well as in obligations relevant for the compliance with copyright law. Any use of copyright protected content requires the authorization of the rightsholder concerned unless relevant copyright exceptions and limitations apply, such as reproductions and extractions of works for the purpose of text and data mining, under certain conditions. But, as rightsholders may choose to reserve their rights over their works to prevent text and data mining, provided the purposes are other than scientific research. In such a case, an authorization needs to be obtained by providers of general-purpose AI models to that end. Along the same vein and in order to increase transparency on the data used in the pre-training and training of general-purpose AI models that may include works or any other subject matter protected by copyright law, providers of these models shall make available a sufficiently detailed summary about the content used for training the model. These obligations, with special regard the two latter ones, are aimed to address the liability concerns more closely associated to copyright infringements and, to a certain extent, privacy, but they seem also to help alleviating the challenge of attributing liability arising from harmful, false or misinforming outputs. Disclosing the content used for pre-training and training might provide insights to identify causes or determining causal links. Potential liability is not totally deactivated though. Nor the complexities to determine and attribute liability.

Transparency presents in the AI Act a second facet that, for a private-law analysis, may have rather enticing and more intriguing implications. It is a disclosure duty or self-revealing feature to identify and mark generated synthetic audio, image, video or text content (Art. 50(2) AI Act). It refers to general-purpose AI systems and is clearly rooted in the generative potential of Generative AI. Thus, providers should implement effective, interoperable, robust and reliable technical solutions to ensure that *the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated.* Some exceptions apply to this obligation where these systems perform assistive function for standard editing or do not substantially alter the input data provided by the deployer or the semantics thereof, or where authorized by law to detect, prevent, investigate or prosecute criminal offences.

Disclosure impacts on perception and consequently, it can reduce the harmful effect, debilitates the reliance as a causal link or simply mitigates the risk of misinterpretation, deceit or frustration of reasonable expectations.

[20] Article 3(64) AI Act:

"high-impact capabilities" means capabilities that match or exceed the capabilities recorded in the most advanced general-purpose AI models.

[21] Article 3(65) AI Act:

"systemic risk" means a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights or the society as a whole, that can be propagated at scale across the value chain.

Various questions are still unsolved. For the purposes of liability, it can be discussed the implications of not marking the artificially generated or manipulated content that may aggravate the harmful consequences of the output. In such cases, total lack of marking may or may not be assimilated to ineffective, non-interoperable or unreliable technical solutions for marking. The end effect may be ultimately similar, but the position of the provider differs on the basis of, among other reasons, whether it is an obligation of result or of best efforts. Yet, in connection with the rebuttable presumption laid down in the draft AILD and the revPLD, the consequences of the infringement of this obligation may require further assessment.

In charting the rules that the AI Act has specifically provided for on general-purpose AI models (or systems), the material scope and the personal scope are and must be treated as closely interwoven. As the AI Act included a separate and distinct concept of general-purpose AI, a specific body of rules emerges strictly associated with these models (or in some cases the general-purpose AI systems as well). Once the "what" is identified, then the "who" is obliged to is immediately resulting from (providers of general-purpose AI models instead of further operators or the relevant operator of the AI supply chain in certain circumstances). The terminological and definitory divide forces to draw a line on the basis of the defined legal terms instead of on grounds of analogous functionality or equivalent effect. That means that in order to assess how the AI Act faces Generative AI, as a paradigm shift in the AI landscape, the rules specifically drafted for general-purpose AI models (and systems) are traced and solely considered. And that has been strictly respected, even if the ultimate definition of AI system, as aligned with the OECD recommendation, has been modified precisely to embrace, among other elements, the content generation capability (OECD, 2024).

From the previous analysis, some findings can be summarized. First, the AI Act has found the way to be permeable to the emergence and the extraordinary popularity gained by Generative AI. To that end, a separate notion of "general-purpose AI (model and system)" was introduced. That drafting decision signals an explicit recognition by the AI Act to the paradigm shift the Generative AI represents and a need to accommodate a regulatory model that had been conceived under a risk-based approach compatible with a general-purpose perspective. Second, the introduction of Generative AI by way of naming it in the AI Act needed to be accompanied by a set of specific provisions suited to the features of Generative AI. So, the AI Act provides for a regulatory model for Generative AI, under a different risk classification, parallel to the use-related risk-based approach that does not work for general-purpose AI models. Nonetheless, this specific set of provisions might not be complete enough to provide a regulatory framework for Generative AI. Third, Generative AI's features as described in previous sections have impact on liability issues. In addition to the necessary adaptations in legacy liability regimes to accommodate AI's distinctive features, as articulated by the revPLD and envisaged by the AILD, additional accommodation might be necessary to face the specific challenges of Generative AI.

Given that, the analysis below of the two legislative actions initiated in the EU to accommodate liability rules to AI challenges is solely focused on appreciating whether the reformed rules, that have been modified to embrace AI systems, have succeeded in incorporating the additional and differential characteristics of Generative AI.

3. Generative AI and liability rules for damages caused by AI systems

The adverse effects resulting from the materialization of AI risks can have a wide variety of manifestations. In certain sectors, significant property damage and personal injury can be anticipated (autonomous vehicles, drones, home automation, assistive robotics). Their applications in financial activities are linked to systemic risks, threats to economic stability and financial integrity or cyclical responses and market shocks. Their use for rankings, recruitment services, content filtering or virtual assistants for complaint management opens the door to a far-reaching debate on their impact on fundamental rights and freedoms – freedom of expression, the right not to be discriminated against,

the right to honor, personality rights – but also on the competitive structure of the market or on the fairness of their practices.

Against such potentially adverse effects, the key question is whether traditional legal regimes are equipped to manage the risks and effectively resolve conflicts arising from such situations in complex technological environments. With the aim of assessing the adequacy of non-contractual liability rules, common distinguishing features, where AI is involved, need to be identified and compared with past and current (non-AI) situations to which the existing rules are (or are assumed to be) well adapted. Such distinctive features – opacity, complexity, vulnerability, openness, data-dependence, autonomy – would allow modelling a legal category of "AI-caused damages" to test the adaptability and adequacy of existing liability rules (Rodríguez de Las Heras Ballell, 2019).

3.1 Dilemmas and legislative policy options in devising an EU response to damages caused by AI systems

The identified distinctive and differential features of AI systems – opacity, complexity, vulnerability, openness, data-dependence, autonomy – prove to have an impact, in some cases of substantial magnitude, on the classical concepts underpinning liability or on the traditional application of liability rules. Given the premise that there is a differential rationale, there are several legislative policy dilemmas for the Union to consider in devising an adequate response to damages caused by AI systems to confront and resolve.

First, to opt for the formulation of specific liability rules for AI or to accommodate the general liability rules to the specificities of AI. Second, to defend a strict liability regime in case of damage caused by AI systems or to maintain a fault-based approach as the main model of fault-based liability. Third, to decide on the level of legal uniformity and harmonization to be achieved at European level for damage caused by AI systems.

The European Parliament Resolution on liability for the operation of AI systems[22] of October 2020, which contained a set of recommendations for a Regulation of the European Parliament and of the Council on civil liability for damage caused by the operation of AI systems, represented a very ambitious position in terms of legislative harmonization and a radical one to solve the three legislative policy dilemmas described above. It consisted of a proposal for a specific liability regime for AI on two levels (strict liability for high-risk systems and fault-based liability for non-high-risk systems) with a high potential for harmonization at European level through the proposed adoption of a Regulation. This proposal did not take its course and has not led to any specific formulation by the Commission in the terms envisaged. Instead, the Commission proposed, using a substantially different approach, the tandem of proposals for a Directive published on 28 September 2022, aimed at revising product liability rules to accommodate "smart products" and (standalone) AI systems, on the one hand (revPLD), and to lighten the burden of proof in fault-based liability actions under national laws on damage caused by AI systems, on the other (AILD).

By departing from the Parliament's 2020 proposal, the Commission takes a clear, and to a certain extent diverging, position on the three policy dilemmas concerning the solution of the AI liability dilemma.

The Commission's approach is less drastic and much less forceful in proposing a specific approach to AI. Despite the revealing name of the proposed Directive – *the Directive on the adaptation of the rules on non-contractual civil liability to artificial intelligence (Artificial Intelligence Liability Directive)* – it is not in fact a Directive establishing liability rules for AI. The objective and expected effect are more modest and certainly more pragmatic and realistic in legislative terms: to establish common rules on disclosure of evidence and burden of proof in non-contractual civil claims for damage caused by an AI system. The modesty of the proposal in the extent (a Directive instead of a Regulation) and

[22]Report with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)), 5 October 2020, at https://www.europarl.europa.eu/doceo/document/A-9-2020-0178_EN.html.

the scope (exclusively rules on disclosure of evidence and burden of proof) is indeed imposed by the low level of harmonization in the EU on civil liability rules that still remain largely national.

Interestingly, the proposed Directive bridges the gap between the AI Act and national fault-based liability laws. While the AI Act, as a regulation, constitutes a decisive attempt to achieve high harmonization at EU level, the rules on fault-based tort liability are essentially national and thus hardly unified. While the AI Act is based on a regulatory approach the AILD aims to close the "redress gap" revealed by the AI Act and to improve the application of regulatory requirements for high-risk AI systems by enhancing their role for the purposes of civil liability claims. In doing so, non-compliance with such requirements triggers the easing of the burden of proof, one of the identified weaknesses of the legacy liability regime that needs to be addressed, by establishing a number of rebuttable presumptions. Unlike Parliament's approach in the 2020 Resolution, the risk-based categorization of the AI system does not lead to strict liability but rather contributes to the rebuttable presumptions. The list-based approach adopted by Parliament's proposal in the 2020 Resolution to categorize high-risk AI systems was not necessarily linked to the AI Act, whereas the AIL consistently bridges the gap by relying on the risk categorization of the AI Act.

As regards the dilemma at the level of harmonization, the Commission, in its proposal of 2021 for an AI Act, renounces achieving maximum harmonization with the adoption of a Regulation on liability rules. Instead, a complex scheme of interactions is devised between the AI Act, the proposed new Directives and national rules on fault-based liability for negligence. The Commission builds two bridges that create an interwoven framework of liability rules for AI systems, although not fully compact at EU level: a bridge between the AI Act and national fault liability rules; and a bridge of complementarity between the product liability regime and fault liability rules.

In this context, the revision of the PLD, that is the second component of the reform duet, plays a key role, with a harmonizing potential that is likely to go beyond the traditional and formal effect of a Directive. By extending the scope of the PLD to include products enabled by AI systems and standalone AI systems, and by adapting some rules of the product liability regime to accommodate the characteristics of such AI systems, the harmonization potential of the PLD is reinforced, extended and leveraged: although this implies proceeding by means of a directive, it is a directive with an express full harmonization clause (Art. 3 revPLD).

The policy strategy adopted by the EU Commission to face the AI liability dilemmas crystallized in two main legislative actions. On the one hand, the profound revision of the product liability regime to accommodate the digital economy, in particular, the use of AI systems, and the circular economy (Recital 3 revPLD). The Commission proposal has been approved by the European Parliament (EP) on 12 March 2024, preceding the subsequent approval of the AI Act on 13 March 2024, and the EP's first reading position has been approved by the Council on 10th October 2024 (PE 7 2024 REV 1). Thus, the EU legal framework for the AI is gaining scale and density. On the other hand, the more modest effort, forced by the EU-Member State competency distribution scheme on the matter, to reach a less ambitious level of harmonization and with less substantial, albeit still relevant and laudable, scope in those rules of national fault-based liability rules that are more bluntly impacted by the AI distinctive features: disclosure of evidence and burden of proof.

3.2 Potentialities and limitations of liability rules for Generative AI

Various features of the proposed (and approved) liability rules for AI systems reveals that, unlike the efforts made to accommodate Generative AI in the regulatory framework (AI Act), further consideration to embrace Generative AI's functional characteristics in the liability narrative is advisable.

3.2.1 Delimiting the scope of application

The first issue to consider is whether Generative AI falls under the scope of application of the AILD and/or revPLD.

On the one hand, due to the strong dependency of the draft AILD on the AI Act (Rodríguez de Las Heras Ballell, 2023), the material scope of the draft rules is demarcated by the notion of "AI system" as defined in the AI Act (Art. 2(1) AILD). As the drafting decision to effectively integrate Generative AI in the logic of the AI Act has been a terminological bifurcation with the addition of "general-purpose AI models" and "general-purpose AI system," the limited referral of the draft AILD to "AI systems" as per the AI Act may invite to be restrictive and literal in the interpretation. In principle, the draft rules would seem not be intended to apply to general-purpose AI models/systems. Nonetheless, there are two relevant points that should not be disregarded. First, a time factor, for the draft AILD, was released before the latest version of the AI Act; therefore, it can be asserted that the text might be plausibly updated, in subsequent steps of the legislative route, and aligned with the approved version of the AI Act. The result of such an update in the material scope of application is uncertain though. Second, a conceptual factor, as any AI system, considering the insights from the explanatory note of the OECD definition, can have generative capabilities (as "content" is explicitly included as a possible output). Nevertheless, the content-generation capability is a feature but, as proposed in this Paper, it is not capturing in its entirety the differential functional characteristics of Generative AI as foundation, general-purpose, multimodal models.

In addition to the limitations inherent to the sole use of the term "AI system," the draft AILD assumes the main risk classification of the AI Act (high-risk, low-risk) that differs from the alternative risk model underpinning the legal regime for general-purpose AI models (standard, systemic risks, free and open source). As it has been pointed out (Novelli et al., 2024a), such a risk classification may fail to take into account the real impact of downstream applications and therefore lead to over-inclusive or under-inclusive risk categories. As a matter of fact, once the general-purpose AI models are not any more considered high-risk by default (as in previous versions) and are subjected to a separate tiered risk classification, the interplay between the AILD and the AI Act seizes.

As for the revPLD, a similar process to unveil its application to Generative AI is to be followed. The key question is whether the enlargement of the scope of application of the PLD through the update of the notion of product manages to embrace Generative AI under its scope.

In deciding to focus on the product liability rules as the preferred route to face and address the challenges posed by AI to liability rules and to use the revision of the Directive as the most realistic way to harmonize AI liability as much as possible, a conceptual dilemma had to be unraveled. Extending the scope of the PLD to include AI systems implies broadening the concept of "product." How far can the concept of "product" be extended without denaturing it? How much strain can the product liability regime be put under to accommodate AI without altering its fundamentals and tenets?

The decision to subject the PLD to a revision by extending its scope was essentially endorsed by the harmonizing potential that such a policy option could obtain. The PLD is highly accepted and applied by industry actors, courts and public authorities. It is settled core of the EU acquis. In a way, by extending the PLD to cover AI systems rather than shifting the focus toward an alternative legislative initiative specific to AI (as the Parliament Resolution of 2020), the Commission is relying on a harmonizing instrument already accepted by industry and widely adopted by national legislators. The revision of the PLD strengthens and broadens the scope of application of EU rules in a harmonized mandate that is already largely uncontested. Fault-based liability rules remains largely national, but Member States have already left the product liability regime to EU rules. Ensuring that this widely accepted model, both by legislators and industry, continues to work well in the digital age should be less contentious than recalibrating the EU-Member States equilibrium in the liability-specific competency landscape.

The decision to accommodate the PLD to AI systems and other challenges of the digital age could not have been achieved simply by an interpretation effort of existing provisions to force them to cover

digital and smart products. Conceptual and practical hurdles had to be overcome. The market needed and expected explicit and clear solutions to ensure legal certainly and improve predictability. Hence, the revision of the PLD entails a process of terminological clarification (of terms such as "product" and "defective") and conceptual recognition of AI systems, as well as the addition of new rules and the incorporation of AI-specific considerations.

Therefore, the revision had to start from the core of the product liability system: the concept of product and the assessment of defectiveness.

Since its adoption in 1985 (Fairgrieve et al., 2016), the PLD has established a definition, for the purposes of the Directive, of a "product"[23] (definition slightly amended in 1999),[24] thereby determining its scope of application. The meaning and scope of this central concept of "product"[25] to cover future technological and market developments is decisive for assessing the versatility of the rules and the full achievement of long-term policy objectives. Moreover, the concept of "product" – its definition, judicial interpretation and application – permeates the entire product liability mechanism and determines, directly or indirectly, the meaning, scope and functioning of its other components (notions such as "defect," "producer" or "manufacturer," as well as issues such as defenses and causation, in addition to the idea of placing a product on the market).

The challenges posed by technological progress and market developments therefore challenge (Rodríguez de Las Heras Ballell, 2020) the core conceptual element of the product liability system but also extend over other elements of the system. The emergence and increasing market penetration of AI-enabled goods, smart products and AI systems are timely catalysts for a profound reflection on the revision (or re-reading) of the PLD through digital lenses. Smart products go beyond the practical and conceptual perimeters of the classical concept of "products" as they were traditionally conceived, devised and constructed in 1985 (European Law Institute, 2021), and even expanded and updated in subsequent years. This inherent limitation of the conceptualization of "product" is despite the clear aspiration of the PLD to produce adaptable standards, and that it because the PLD transpired a rather recognizable industrial and post-industrial flavor.

The stylized definition of "product" in its original wording (Article 2 PLD) covers all movable property, including property incorporated in, affixed to or associated with other movable or immovable property. It is a broad and malleable definition that effectively covers the diverse typology of products derived from industry and market innovation. It is indeed comprehensive and reasonably future-proof, but it is still somewhat constrained by an underlying industrial logic. Thus, smart products, AI systems and even software call into question not only the concept of "product" itself in the PLD but also and fundamentally the distinction between products and services (Hojnik, 2017), between assets and data (Gemignani, 1980; Green & Saidov, 2007; Stapleton, 1989; Weber, 2012) and between objects and subjects in the modern economy, and thus, in the formulation and application of legal rules as well. This is precisely why smart products invite a bold revision of the conceptual basis and policy vectors underpinning the PLD.

[23] Article 2 PLD:

> For the purposes of this Directive, "product" shall mean any movable property, with the exception of primary agricultural products and game, even if incorporated in other movable property or in immovable property. Primary agricultural products' means products of the soil, of stock farming and of fisheries, excluding products which have undergone primary processing. "Product" includes electricity.

[24] Directive 1999/34/EC of the European Parliament and of the Council of 10.5.1999 amending Council Directive 85/374/EEC on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 141, 4.6.1999, pp. 20–21. The amendment consisted in deleting the exclusion of primary agricultural products and hunting. Subsequently, the scope was explicitly extended by amending the definition in 1999.

[25] For the purposes of this Directive, "'product' means any movable property, even if it is incorporated into other movable property or into immovable property. The term 'product' includes electricity."

AI-enabled products challenge the legacy system of product liability because they blur the line between products and services (Vandermerwe & Rada, 1998; Araujo & Spring, 2006; Gadrey, 2000; Hill, 1999; Parry et al., 2011; Rathmell, 1966), blur the contours of products as single units by transforming them into complex ecosystems, that evolve throughout their lifecycle by way of updating and upgrading, that are fed by data and interact with the environment as if they were, metaphorically, "*quasi-living* beings" (Rodríguez de Las Heras Ballell, 2006). These are the disruptive aspects of AI-enabled products that perfectly capture the specific characteristics of AI systems, as presented above (*supra* II).

Accordingly, the decision of the revised PLD to amend the definition of "product" and to explicitly include "software" therein[26] is crucial and to be welcomed. But even more telling is the express clarification in the explanatory note (p. 5) of the first proposal of the Commission[27] and Recital 13 of the revPLD[28] that AI systems and products based on AI are "products" for the purposes of the PLD.

Accordingly, providers of AI systems (according to the AI Act)[29] will be considered manufacturers. This strong drafting solution, by explicitly mentioning "software" in the definition of product, provides significant clarification, although, unfortunately, as with any drafting change, some uncertainties are alleviated while new ones are created. Thus, as the European Law Institute pointed out (Koch et al., 2022), it remains unclear whether other digital content that may be functionally equivalent to software is included as a product despite not performing specific tasks on its own, as well as whether SaaS (software as a service) is also included or not, provided that from the victim's perspective no distinction in the marketing model (provided as a stand-alone product or under a subscription contract) is relevant when it comes to securing compensation. Clarification in that sense has been added in the recitals, confirming that, for the purposes of the PLD, software is a product irrespective of the mode of its supply or usage, including SaaS models (Recital 13 revPLD).

Beyond the broadening of the meaning given to the concept of "product," the amended definition of "component" is certainly much more revolutionary and tempting. According to the revised text (Art. 4(4) revPLD), "component part" means any item, whether tangible or intangible, or raw material or any related service, that is integrated into, or inter-connected with, a product. This definition, along with the following definition of "related service" (Art. 4(3) revPLD) – "means a digital service that is integrated into, or inter-connected with, a product in such a way that its absence would prevent

[26] Art. 4(1) revPLD:

> "product" means all movables, even if integrated into, or inter-connected with, another movable or an immovable; it includes electricity, digital manufacturing files, raw materials and software.

[27] COM(2022) 495 – Proposal for a directive of the European Parliament and of the Council on liability for defective products:

> in respect of AI in particular, this proposal confirms that AI systems and AI-enabled goods are "products" and therefore fall within the PLD's scope, meaning that compensation is available when defective AI causes damage, without the injured person having to prove the manufacturer's fault, just like for any other product.

[28] Recital 13 revPLD:

> Products in the digital age can be tangible or intangible. Software, such as operating systems, firmware, computer programs, applications or AI systems, is increasingly common on the market and plays an increasingly important role for product safety. Software is capable of being placed on the market as a standalone product or can subsequently be integrated into other products as a component, and it is capable of causing damage through its execution.

[29] Art.3 (3) AI Act:

> "provider" means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge.

the product from performing one or more of its functions" – is full of elements that emerge from the AI-driven paradigm shift: interconnection, integration, performance and fulfilment of expected functions and, of course, related service. Services are conquering the terrain of "products." And this conquest is visible, explicit and revealing of the progressive dilution of conceptual boundaries.

The categorization of "related service" as a component for the purposes of product liability rules is also embodying the definitive irruption of smart products in the market and the necessary adaption of the traditional liability rules to the peculiar characteristics of such classes of products. While the revised product liability regime (as stresses in Recital 17 revPLD) persists in distinguishing products and services and excluding the latter from its scope of application, it cannot ignore how new products operate. The safety of smart products largely depends upon integrated, embedded or interconnected digital services as much as they do upon physical components. Therefore, such "related services" must be considered components for the purposes of product liability; otherwise, the consistency of the liability rules would be undermined and the coherence of the regime would be fractured.

Once the then indelible line between products and services is, to a certain extent, erased, a constellation of scenarios opens up that required a rather subtle and case-by-case analysis. Thus, data supply services or monitoring services are covered by the modernized notion of "component" (such as the continuous supply of traffic data in a navigation system, or a temperature control service that monitors and regulates certain functions of a smart home system). More nuanced has to be the categorization of the general internet access services. In principle, these services are not "related services" as they are not under the manufacturer's control. Nevertheless, as explained in Recital 17 revPLD, should a product that relies on internet access services fail to maintain safety in case of loss of connectivity, the liability rules may find such product defective.

The drafting option taken in the revision of the PLD provides less conclusive insights about the intended or actual coverage of Generative AI by the product liability regime. The wording of the scoping provision is not straightforward and the interpretative aids of the recitals are generally referring to "AI systems." Again, the terminological bifurcation of the AI Act adds complexity to the debate (general-purpose AI models or systems are not covered under the primary notion of "AI system").

Therefore, a functional approach is more advisable. Should the AI Act define general-purpose AI models as essential components of downstream AI systems, at least two possible scenarios can be envisioned. On the one hand, the provider of the general-purpose AI model is deemed as the manufacturer of a component (Art. 8(1) revPLD) that turns out to be defective. On the other hand, as the general-purpose AI models is to be fine-tuned and adapted to downstream applications, the notion of "substantial modification" may apply (Art. 8(2) revPLD). The downstream fine-tuning process might be read as a "refurbishing" process. Hence, the provider of the AI system integrating and adapting the general-purpose AI models becomes a manufacturer for the purposes of the revPLD.

3.2.2 Expected use and presumptions

Beyond the unresolved question of the scope of application, in delving into the liability rules as set out both in the AILD and in the revPLD, other points of inconsistency or friction surface.

In the revPLD, along with the notion of product, the concept of "defectiveness" is pivotal. The revised wording of Article 7 revPLD on defectiveness clearly shows that attention is paid to the specificities of smart products and AI systems. Defectiveness is defined in relation to *the safety that a person is entitled to expect or that is required by the applicable law.* Thus, defectiveness is gauged against safety expectations. It is an objective assessment of the safety that the public at large is entitled to expected, on the basis of certain factors. The very special features of AI systems and AI-driven products do precisely impact on those factors to be taking into account in the assessment of the safety expectations.

The solution implemented in the revision of the PLD has been to extend the list of factors to be taken into account in the assessment of the defectiveness of a product and to include those more effectively acknowledging the distinctive and differential features of AI systems and smart products. The new factors included in the list significantly enlarge the previous wording of the former (shorter)

Article 6 of the PLD[30] before the revision. The selected new factors transpire the distinctive features of AI. This is particularly revealing in, *inter alia*, (a) (…) the instructions for its assembly, installation, use and maintenance; (c) the effect on the product of any ability to continue to learn or acquire new features after it is placed on the market or put into service; (d) the effect on the product of other products that can reasonably be expected to be used together with the product, including by means of inter-connection; or (e) relevant security requirements of the product, including security-relevant cybersecurity requirements.

While the drafting solution is well received and seems to succeed in improving clarity and providing guidance in the assessment of defectiveness with regard to AI-based products, it does not seem suited to Generative AI. The generality of these models sets a large distance with "reasonably foreseeable use" of the product (Art. 7(2)b revPLD). Even the idea of "reasonable safety expectations" as the yardstick to assess defectiveness loses efficacy in the context of general-purpose AI models, as it might not be easily anticipated such expectations due to the generality factor.

Therefore, it is advocated (Novelli et al., 2024b) that the assessment of defectiveness (for the revPLD) and fault (for the AILD) should be rather focused on monitoring measures that prioritize oversight and supervision of random or unexpected outputs instead of more "static" compliance requirements.

Yet, both the draft AILD and the revPLD spotlight the need to face the asymmetry and complexities that AI's features allocate on the victim. It has been convincingly argued (European Commission, 2019) that the opacity, autonomy and complexity of AI systems call into question the balance of interests underlying the current distribution of the burden of proof in the former PLD (Article 4)[31] as well as in any claim for damages. The injured party will face significant difficulties in proving a defect and the causal link between a defect and the damage in the face of very opaque and complex decision-making based on AI. Mere transparency ("opening the black box") may not succeed in improving the victim's position due to the complexity of the underlying algorithmic bases if these are not clarified by an effective explanation of the reasons, the trajectory of decision-making and possible critical deviations or biases. Similarly, efforts to gather relevant evidence, trace actions along the causal chain and collect data may be fruitless, deterrent or unaffordable without the cooperation of the actors involved in the functioning of the technology ecosystem and along the AI supply chain.

With the proposal for the AILD, the Commission intends to address one of the most visible sticking points in accommodating damage caused by AI systems in traditional fault-based liability rules, namely the disclosure of evidence and the burden of proof. The product liability model required similar modification to rebalance the resulting asymmetry as well.

Recognizing the asymmetry that the disruptive features described above can cause to the detriment of the injured party, the legal logic underpinning the product liability regime must be reviewed. Shifting, alleviating or lowering the standard of proof in favor of the weaker party are reasonable responses to remedy the imbalance.

[30] Former Article 6 PLD:

1. A product is defective when it does not provide the safety which a person is entitled to expect, taking all circumstances into account, including:
   (a) the presentation of the product;
   (b) the use to which it could reasonably be expected that the product would be put;
   (c) the time when the product was put into circulation.
2. A product shall not be considered defective for the sole reason that a better product is subsequently put into circulation.

[31] Article 4 PLD before the revision:

The injured person shall be required to prove the damage, the defect and the causal relationship between defect and damage.

Articles 9 and 10 revPLD combine two of the methods aimed at facilitating proof by the victim: rules on disclosure of evidence (Art. 9) and rebuttable presumptions (Art. 10). Thus, the PLD also benefits from techniques aimed at easing the burden of proof such as those incorporated in the proposed AILD. While the burden to prove the defectiveness of the product, the damage suffered and the causal link is still on the claimant, where certain conditions are met, defectiveness, causal link or both shall be presumed. Thus, the burden of proof is alleviated where technical or scientific complexity, or opacity render the evidential efforts of the victim excessive, or unreasonable and may lead to an unfair apportionment of risk.

These two requirements are not poised to seamlessly apply to Generative AI.

First, as there are specific references to high-risk systems what does not match with the tiered risk classification applicable to general-purpose AI models. As explained above, the high-risk category belongs to the risk classification that the AI Act deploys for AI systems, but general-purpose AI models do not (and cannot) be subjected to such a risk classification as risk levels depend upon the intended use. In a strict interpretation, it could be argued that as general-purpose AI models cannot be categorized as high-risk systems, any liability rule based on or referring to provisions applicable to high-risk systems would simply become inapplicable.

Second, as the available information to be disclosed by general-purpose AI models depends upon their own risk classification (standard, systemic risk, free and open-license), the content of the disclosure is not homogeneous and varies. Therefore, even in the case that they are not deemed incompatible, they are simply unsuited.

Third, as the AILD misaligns with the AI Act in respect of the disclosure obligations that are relevant for the presumption but that are not applicable to general-purpose AI models in the final version of the text. It may be considered as a mere temporary mismatch, which might be readjusted if the AILD goes forward. Nevertheless, it reveals a profounder misalignment. Despite the significant efforts made to accommodate Generative AI in the various initiatives adopted or in motion aimed to address AI-related issues, Generative AI proves to represent a paradigm shift in the AI landscape. Then, not surprisingly, the texts envisioned and drafted in a pre-Generative AI context show inconsistencies.

Consequently, the terminological accommodation and the relative substantive incorporation of Generative AI in the AI Act do not lead to affirm that a complete and future-proof framework for Generative AI has been established. Interpretation by analogy is not convincing in the facing of the paradigm-shifting character of Generative AI. Thus, presuming that the revPLD provisions would properly work for general-purpose AI models has no sufficient base. Future work and closer consideration are advisable and necessary.

4. Final remarks and future work

The advent of Generative AI has profoundly and substantially impacted on the regulatory process on AI in the EU. As Generative AI embodies a paradigm shift in the AI scene, legal concepts, rules and regulatory strategy need to be revised and adapted accordingly. The AI Act as well as the two legal actions aimed at accommodating liability rules to AI systems (draft AILD and revPLD) have been primarily conceived, developed and implemented to address the distinctive features of AI systems. While Generative AI is, nevertheless, neither ignored nor disregarded in the final version of the AI Act, their differential, functional features may lead to certain misalignment with some regulatory assumptions and reveal various points of frictions. As a matter of fact, the European AI Office launched a multi-stakeholder consultation on trustworthy general-purpose AI models under the AI Act and a call for the drafting of the first General-Purpose AI Code of Practice.[32] The kick-off of the process for drawing-up of the first Code of Practice for general-purpose AI models took place on 30 September 2024. Endorsing the finding that AI Act provisions are not sufficiently clear and not always

[32]https://digital-strategy.ec.europa.eu/en/consultations/ai-act-have-your-say-trustworthy-general-purpose-ai

suited to Generative AI, this Code of Practice aims to facilitate the proper application of the AI Act's rules for general-purpose AI models, including transparency and copyright-related rules, systemic risk taxonomy, risk assessment and mitigation measures.

Under an innovative regulatory strategy, AI Act seems to act as a framework regulation that needs to be concretized in some points, accompanied by guidance to implementation or complemented by codes of practice in order to made it work in a dynamic environment without being fossilized. The first conclusion then is that future work to make the AI Act work for Generative AI is necessary.

The Paper has traced the drafting process aimed to incorporate Generative AI in the AI Act with different terminological and conceptual solutions. The finally adopted bifurcation between "AI systems" and "general-purpose AI models (and systems)" crystallizes a policy decision that clearly influences the applicable regime as well as the liability rules. After coping with the inherent clash of the generality factor of Generative AI with a purpose-specific risk-based approach, the AI Act has separated the risk classification into two models. Hence, the general-purpose AI models are integrated into the AI Act framework but under an alternative risk classification and subject to a body of selected rules. That impacts on the liability regimes that directly or indirectly cling from the AI Act.

Besides, general-purpose AI models while serving as "foundational models" play a critical role in the AI supply chain as they are a key component. Thus, they become concurrently "points of reinforcement" and "points of failure." Both strengths and weaknesses of general-purpose AI models amplify and pervade along and all over the AI supply chain. That multiplying factor is to be considered in the liability assessment and in the regulatory approach.

The purpose versality of general-purpose AI models, even in performing untrained or unexpected tasks, opens up a wide range of possible liability scenarios. The AI Act neither provides a complete legal regime for Generative AI nor identifies all risk scenarios. On the contrary, as Generative AI permeates economic and social activities, risks diversify and liability cases escalate. From criminal cases to contractual liability situations, from privacy or copyright infringement to various fundamental rights' violation, the generality, multimodality and foundation character of general-purpose AI models deploy a chain of infringements of highly diverse nature.

Therefore, in order to map the possible liability risks Generative AI may bring about, it is rather advisable to identify the critical points and outline a plan or strategy to conduct the assessment instead of trying to envision all imaginable scenarios. The assessment needs to start by making a distinction between three different stages: training, use or operation and generation of outputs. Training exposes the models to more intense risks of privacy and copyright infringements, while the use of the model for a specific purpose is going to open up a wide array of differing situations in terms of risk, liability and regulatory compliance. In the case of general-purpose AI models, this distinction is more distinguishable precisely due to the level of generality and the need of subsequent fine-tuning and adaptation. As explained above, the AI Act situates at the first level and allocates on the providers of these models under certain conditions (as per the tiered risk classification) certain compliance requirements aimed to address the most visible risks (privacy, or copyright infringement, in particular). The second level, where the purpose is identified, leads to the core of the AI Act with a battery of measures and obligations applicable to the specific uses. Yet, on the third level, a constellation of possible liability scenarios deploys. Generated outputs can raise issues related to misinformation, hatred speech, discrimination, copyright infringement, criminal offences of various types, misleading precontractual messages or, *inter alia*, unfair advertising. Potential risks are uncountable.

The Paper has analyzed in depth the revPLD and the AILD to test their adequacy and suitability for Generative AI. Literal, functional and even teleological interpretation leads to hesitant conclusions. The revised liability rules of the defective product liability regime and the proposed new rules for fault-based non-contractual liability do not seem to fully embrace Generative AI. While terminological misalignment can be solved or even can be considered surmountable with a generous interpretation, the functional mismatch reveals that a more profound reconsideration is needed.

Hence, future work should be guided by a macro-analysis aimed to identify and formulate effective and predictable attribution rules to allocate risks and liability in Generative AI cases along the AI supply chain, from the provider to the user, that may serve as an analytical framework for any further liability scenario.

Acknowledgments. Funding for APC: Universidad Carlos III de Madrid (Agreement CRUE-Madroño 2024).

Competing interest. The author declares none.

References

Anderljung, M., Barnhart, J., Korinek, A., Leung, J., O'Keefe, C., Whittlestone, J., Avin, S., Brundage, M., Bullock, J., Cass-Beggs, D., Chang, B., Collins, T., Fist, T., Hadfield, G., Hayes, A., Ho, L., Hooker, S., Horvitz, E., Kolt, N., Schuett, J., Shavit, Y., Siddarth, D., Trager, R., & Wolf, K. (2023). Frontier AI regulation: Managing emerging risks to public safety. arXiv preprint arXiv:2307.03718.

Araujo, L., & Spring, M. (2006). Services, products and the institutional structure of production. *Industrial Marketing Management*, *35*(7), 797–805.

Boden, M. A. (2009). Computer models of creativity. *AI Magazine*, *30*(3), 23–34. https://doi.org/10.1609/aimag.v30i3.2254

Bommasani, R, Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg. J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Quincy Davis, J., Demszky, D., Donahue, C., Doumbouya, M., Durmus, E., Ermon, S., Etchemendy, J., Ethayarajh, K., Fei-Fei, L., Finn, C., Gale, T., Gillespie, L., Goel, K., Goodman, N., Grossman, S., Guha, N., Hashimoto, T., Henderson, P., Hewitt, J., Ho, D.E., Hong, J., Hsu, K., Huang, J., Icard, T., Jain, S., Jurafsky, D., Kalluri, P., Karamcheti, S., Keeling, G., Khani, F., Khattab, O., Wei Koh, P., Krass, M., Krishna, R., Kuditipudi, R., Kumar, A., Ladhak, F., Lee, M., Lee, T., Leskovec, J., Levent, I., Li, X.L., Li, X., Ma, T., Malik, A., Manning, C.D., Mirchandani, S., Mitchell, E., Munyikwa, Z., Nair, S., Narayan, A., Narayanan, D., Newman, B., Nie, A., Niebles, J.C., Nilforoshan, H., Nyarko, J., Ogut, G., Orr, L., Papadimitriou, I., Sung, Park, Piech, C., Portelance, E., Potts, C., Raghunathan, A., Reich, R., Ren, H., Rong, F., Roohani, Y., Ruiz, C., Ryan, J., Ré, C., Sadigh, D., Sagawa, S., Santhanam, K., Shih, A., Srinivasan, K., Tamkin, A., Taori, R., Thomas, A.W., Tramèr, F., Wang, R.E., Wang, W., Wu, B., Wu, J., Wu, Y., Xie, S.M., Yasunaga, M., You, J., Zaharia, M., Zhang, M., Zhang, T., Zhang, X., Zhang, Y., Zheng, L., Zhou, K., & Liang, P. (2021). On the opportunities and risks of foundation models (manuscript), https://arxiv.org/abs/2108.07258

Bozinovski, S. (2020). Reminder of the first paper on transfer learning in neural networks, 1976. *Informatica*, *44*(3), 291–302.

Chan, A., Salganik, R., Markelius, A., Pang, C., Rajkumar, N., Krasheninnikov, D., Langosco, L., He, Z., Duan, Y., Carroll, M., Lin, M., Mayhew, A., Collins, K., Molamohammadi, M., Burden, J., Zhao, W., Rismani, S., Voudouris, K., Bhatt, U., Weller, A., Krueger, D., and Maharaj, T. (2023). *Harms from increasingly agentic algorithmic systems*. IEEE Computer Society, March 2022, https://arxiv.org/abs/2302.10329.

European Commission. (2020, February 19). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions. *Shaping Europe's Digital Future*. COM/2020/67 final, Brussels.

European Commission. (2021 December). *Behavioural study on the link between challenges of artificial intelligence for Member States' civil liability rules and consumer attitudes towards AI-enabled products and services*. JUST/2020/RCON/FW/CIVI/0065.

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

European Law Institute. (2021). Guiding principles for updating the product liability directive for the digital age (ELI Innovation Paper Series) https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Guiding_Principles_for_Updating_the_PLD_for_the_Digital_Age.pdf.

Fairgrieve, D., Howells, G., Mogelvang-Hansen, P., Streatmans, G., Verhoeven, D., Machnikowski, P., Janssen, A., and Schulze, R. (2016). Product liability directive. In P. Machnikowski (Ed.), *European product liability: An analysis of the state of the art in the era of new technologies. Principles of European Tort Law* (pp. 17–108). Intersentia.

Gadrey, J. (2000). The characterization of goods and services: An alternative approach. *Review of Income and Wealth*, *46*(3), 369–387.

Gemignani, M. (1980). Product liability and software. *Rutgers Computer & Technology Law Journal*, *8*, 173.

Green, S., & Saidov, D. (2007). Software as goods. *Journal of Business Law*, *161*, 161–181.

Grimmelmann, J. (2015). Copyright for literate robots. *Iowa Law Review*, *101*, 657–681.

Hacker, P. (2021). A legal framework for AI training data—From first principles to the Artificial Intelligence Act. *Law, Innovation and Technology*, *13*(2), 257–301. https://doi.org/10.1080/17579961.2021.1977219

Hacker, P. (2023a). The European AI liability directives–critique of a half-hearted approach and lessons for the future. *Computer Law & Security Review*, *51*, 105871. https://doi.org/10.1016/j.clsr.2023.105871

Hacker, P. (2023b, December). What's missing from the EU AI Act: Addressing the four key challenges of large language models. *Verfassungsblog*.

Hacker, P., Engel, A., & Mauer, M. (2023). *Regulating ChatGPT and other large generative AI models* [Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)] (pp. 1112–1123). https://doi.org/10.1145/3593013.3594067.

Hill, P. (1999). Tangibles, intangibles and services: A new taxonomy for the classification of output. *The Canadian Journal of Economics / Revue Canadienne d'Economique*, *32*(2), 426–446.

Hojnik, J. (2017). Technology neutral EU law: Digital goods within the traditional goods/services distinction. *International Journal of Law and Information Technology*, *25*(1), 63–84.

Koch, B., Borghetti, J.-S., & Machnikowski, P. (2022). Response of the European law institute to the public consultation on civil liability – Adapting liability rules to the digital age and artificial intelligence. *Journal of European Tort Law*, *13*(1), 25–63.

Nerlich, B. (2023). *Frontier AI: Tracing the origin of a concept*. https://blogs.nottingham.ac.uk/makingsciencepublic/2023/10/20/frontier-ai-tracing-the-origin-of-a-concept/.

NGO, R., Chan, L., & Mindermann, S. (2023). *The alignment problem from a deep learning perspective*. *arXiv*: 2209.00626 [cs.AI].

Novelli, C., Casolari, F., Hacker, P., Spedicate, G., & Floridi, L. (2024b, January 14). Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity. Available at SSRN: https://ssrn.com/abstract=4694565 or https://dx.doi.org/10.2139/ssrn.4694565.

Novelli, C., Casolari, F., & Rotolo, A. (2024a). AI risk assessment: A scenario-based, proportional methodology for the AI Act. *Digital Society*, *3*(1). https://dx.doi.org/10.1007/s44206-024-00095-1

OECD. (2019a). AI terms & concepts. https://oecd.ai/en/ai-principles. accessed 25 November 2024.

OECD. (2019b, May). *Recommendation of the council on artificial intelligence*. OECD/LEGAL/0449.

OECD. (2023a). *Recommendation of the council on artificial intelligence*, OECD/LEGAL/0449. Revision 8 November 2023.

OECD. (2023b). Initial policy considerations for generative artificial intelligence. *OECD Artificial Intelligence Papers*, 1.

OECD. (2024). Explanatory memorandum on the updated OECD definition of an AI system. *OECD Artificial Intelligence Papers*, *8*.

Parry, G., Newnes, L., & Xiaoxi, H. (2011). Goods, products and services. In M. Macintyre, et al. (Ed.), *Service design and delivery* (pp. 19–29). Springer.

Perez, E., Ringer, S., Lukosiute, K., Nguyen, K., Chen, E., Heiner, S., Pettit, C., Olsson, C., Kundu, S., Kadavath, S., Jones, A., Chen, A., Mann, B.,Israel, B., Seethor, B., Mckinnon, C., Olah, C., Yan, D., Amodei, D., Drain, D., Li, D., Tran.Johson, E., Khundadzem, G., Kernion, J., Landis, J., Kerr, J., Mueller, J., Hyun, J., Landau, J., Ndousse, K., Galdberg, L., Lovitt, L., Lucas, M., Sellitto, M., Zhang, M., Kingsland, N., Elhage, N., Joseph, N., Mercado, N., Dassarma, N., Rauschm, O., Larson, R., Mccandlish, S., Johston, S., Kravec, S., El Showk, S., Lanham, T., Telleen-Lawton, T., Brown, T., Henigham, T., Hume, T., Bai, Y., Hatfield-Dodds, Z., Clark, J., Browman, S.R., Askell, A., Grosse, R., Hernandez, D., Ganguli, D., Hubinger, E., Schiefer, N., & Kaplan, J. (2023). Discovering language model behaviors with model-written evaluations. *Findings of the Association for Computational Linguistics: ACL*, 13387.

Pratt, L.Y., Mostow, J., & Kamm, C.A. (1991). Direct transfer of learned information among neural networks. In *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI–91)* (pp. 584–589. Anaheim, CA.

Raji, I.D., Kumar, E., Horowitz, A., & Selbst, A.D.,(2022). The fallacy of AI functionality. *FAccT* 22, 959.

Rathmell, J. M. (1966). What is meant by services? *Journal of Marketing*, *30*(4), 32–36.

Rodríguez de Las Heras Ballell, T. (2006). La responsabilidad por software defectuoso en la contratación mercantil. *Revista Aranzadi de Derecho Y Nuevas Tecnologías*, *10*, 83–110.

Rodríguez de Las Heras Ballell, T. (2019). Legal challenges of artificial intelligence: Modelling the disruptive features of emerging technologies and assessing their possible legal impact. *Uniform Law Review*, *1*(2019), 1–13.

Rodríguez de Las Heras Ballell, T. (2020). Embracing technological disruption in international transactions: Challenges for legal harmonization. In Benicke, C., and Huber, S (Eds.), *National, international, trasnational: Harmonischer Dreiklang Im Recht, Festschrift Für Herbert Kronke, Zum. 70 Geburstag, Bielefeld*. (pp. 1223–1233). Giesekin Verlag.

Rodríguez de Las Heras Ballell, T. (2023). The revision of the product liability directive: A key piece in the artificial intelligence liability puzzle. *ERA Forum*, *24*(2), 247–259. doi:10.1007/s12027-023-00751-y

Sartor, G., Lagioia, F., & Contissa, G. (2018, October 11). The use of copyrighted works by AI systems: Art works in the data mill. *SSRN Electronic Journal*. Available at SSRN: https://ssrn.com/abstract=3264742 or.

Stapleton, J. (1989). Software, information and the concept of product. *Tel Aviv University Studies in Law*, *9*, 147–164.

Vandermerwe, S., & Rada, J. (1998). Servitization of business: Adding value by adding services. *European Management Journal*, *6*(4), 314–324.

Volokh, E., Decio, A., & Formenti, L. (2023). Large libel models? Liability for AI output. J. free speech L. *Cancer Research Communications*, *3*, 489–558.

Weber, L. A. (2012). Bad bytes: The application of strict products liability to computer software. *St. John's Law Review*, *66*(2), 469–486.

World Economic Forum (WEF). (2018). *Fourth industrial revolution for the Earth series harnessing artificial intelligence for the Earth*.

Ye, H., Liu, T., Zhang, A., Hua, W., & Jia, W. (2023). Cognitive mirage: A review of Hallucinations in large language models. arXiv. https://doi.org/10.48550/arXiv.2309.06794.

Full Professor of Commercial Law, Universidad Carlos III de Madrid. Member of the Austrian Academy of Sciences. She is Spain's delegate and expert to UNCITRAL and UNIDROIT in various projects; and member of European Commission Expert Groups on New Technologies and Liability; on B2B Data Sharing; and on Platforms Economy. Member of ELI Council and ELI Executive Committee. She was awarded a scholarship from the European Central Bank Legal Programme in 2018 on Fintech Regulation and Sir Roy Goode Scholar at UNIDROIT, 2021–2022. She is an Arbitrator for the Madrid Arbitration Court. She held visiting professor positions in various universities, *inter alia*, Visiting Professor at the National University of Singapore, Visiting Scientist at the University of Turin, the Chair of Excellence (Santander-UC3M) at Oxford University, Distinguished Law Professor at Tulane Law School.

Cite this article: Rodríguez de Las Heras Ballell T. (2025). Mapping Generative AI rules and liability scenarios in the AI Act, and in the proposed EU liability rules for AI liability. *Cambridge Forum on AI: Law and Governance* **1**, e5, 1–23. https://doi.org/10.1017/cfl.2024.8