Computer Law & Security Review 53 (2024) 105984

Contents lists available at ScienceDirect

Computer Law & Security Review: The International Journal of Technology Law and Practice

journal homepage: www.elsevier.com/locate/clsr

The EU Regulatory approach(es) to AI liability, and its Application to the financial services market

Maria Lillà Montagnani [a,#,*], Marie-Claire Najjar [b,#], Antonio Davola [c,#]

[a] Commercial Law at Bocconi University and director of the Bocconi LL.M in European Business and Social Law, Italy
[b] Bocconi University and co-coordinator of the Bocconi LL.M. in European Business and Social Law, Italy
[c] Economic Law at the University of Bari 'Aldo Moro', Italy

ARTICLE INFO

Keywords:
AI governance
AI liability
AI Act
AI liability directive

ABSTRACT

The continued progress of Artificial Intelligence (AI) can benefit different aspects of society and various fields of the economy, yet pose crucial risks to both those who offer such technologies and those who use them. These risks are emphasized by the unpredictability of developments in AI technology (such as the increased level of autonomy of self-learning systems), which renders it even more difficult to build a comprehensive legal framework accounting for all potential legal and ethical issues arising from the use of AI. As such, enforcement authorities are facing increased difficulties in checking compliance with applicable legislation and assessing liability, due to the specific features of AI, – namely: complexity, opacity, autonomy, unpredictability, openness, data-drivenness, and vulnerability. These problems are particularly significant in areas, such as financial markets, in which consequences arising from malfunctioning of AI systems are likely to have a major impact both in terms of individuals' protection, and of overall market stability. This scenario challenges policymaking in an increasingly digital and global context, where it becomes difficult for regulators to predict and face the impact of AI systems on economy and society, to make sure that they are human-centric, ethical, explainable, sustainable and respectful of fundamental rights and values. The European Union has been dedicating increased attention to filling the gap between the existing legal framework and AI. Some of the legislative proposals in consideration call for preventive legislation and introduce obligations on different actors – such as the AI Act – while others have a compensatory scope and seek to build a liability framework – such as the proposed AI Liability Directive and revised Product Liability Directive. At the same time, cross-sectorial regulations shall coexist with sector-specific initiatives, and the rules they establish. The present paper starts by assessing the fit of the existing European liability regime(s) with the constantly evolving AI landscape, by identifying the normative foundations on which a liability regime for such technology should be built on. It then addresses the proposed additions and revisions to the legislation, focusing on how they seek to govern AI systems, with a major focus on their implications on highly-regulated complex systems such as financial markets. Finally, it considers potential additional measures that could continue to strike a balance between the interests of all parties, namely by seeking to reduce the inherent risks that accompany the use of AI and to leverage its major benefits for our society and economy.

1. Introduction

In 2017, Monaco-based investment manager Tyndaris SAM signed an agreement with VWM Limited to manage an account using an artificial intelligence (AI) system for investment decisions. The AI system, known as the K1 supercomputer, was capable of predicting market sentiment and providing trading signals based on real-time news and social media data. It had been extensively back-tested and live-tested before trading

* Corresponding author.
E-mail address: lilla.montagnani@unibocconi.it (M.L. Montagnani).

[#] The present paper is the joint effort of all the Authors, who contributed equally to its development. In particular, Maria Lillà Montagnani contributed mostly to Sects 1, 2.1, 2.2, Marie-Claire Najjar contributed mostly to Sects. 3.1, 3.2, and Antonio Davola contributed mostly to Sects. 2.3 and 3.3

https://doi.org/10.1016/j.clsr.2024.105984

Available online 31 May 2024

began in December 2017.[1] However, VWM quickly suffered significant losses, amounting to approximately US $22 million overall. In response, VWM wrote to Tyndaris demanding that trading be suspended until further notice. Tyndaris claimed approximately US $3 million from VWM in unpaid fees (for the account management) and eventually commenced proceedings in the English High Court. VWM counterclaimed, seeking to recover its losses on the basis that it had invested in the managed account in reliance on misrepresentations by Tyndaris regarding the capabilities of the K1 supercomputer.[2]

This case – which unfortunately for the development of tort law in this important area was settled – highlights the central role that AI[3] plays in financial markets and in general in contemporary markets, and the risks associated with its deployment when not properly supervised. AI systems develop in a way that makes them pursue their tasks with diverse degrees of autonomy.[4] Their new and enhanced potential brings in risks, or increases the existing ones, for both those who offer them and those who use them. Indeed, such technologies may have unintended effects or be used for malicious purposes. They can lead not only to discrimination and biases,[5] but also to violation of IP and personality rights,[6] unauthorized access and cybersecurity vulnerabilities,[7] and errors that can ruin a persons life.[8] As a matter of fact, while AI can bring benefits such as increased productivity and can transform products, services, activities, procedures, and practices in several economic sectors and in relation to many aspects of society – such as health,[9] sustainability,[10] sports,[11] transportation[12] – it can also lead to unintended effects, biases, and violations of fundamental rights.

The double-edged nature of AI raises challenges for regulators and policymakers who need to balance its potential benefits with its potential harms. In regulating AI, like for other technologies, the debate about technology neutrality also remains a central issue. The principle of technology neutrality holds that regulations and laws should neither privilege nor penalise one technology (or set thereof) over another. More generally, there is a lively debate around the ethical issues raised by new technologies, as analysed in the UNESCO's Recommendation on the Ethics of Artificial Intelligence.[13] The regulation of AI in a way that strikes a balance between protection and innovation requires a 'responsible innovation'[14] strategy that ensures that AI is human-centric, ethical, explainable, sustainable, and respectful of fundamental rights and values. To create an environment of trust and accountability, legal rules on civil liability must be designed to address

[1] M Tanna and W Dunning, 'Who (if anyone) is liable when an artificial intelligence (AI)-powered trading / investment system causes substantial losses for an investor?' (*Simmons-Simmons*, 13 November 2019) <https://www.simmons-simmons.com/en/publications/ck2xifd2ddmrq0b48u46j2nns/ai-powered-investments-who-if-anyone-is-liable-when-it-goes-wrong-tyndaris-v-vwm>.

[2] J Kahn, 'Why do so few businesses see financial gains from using A.I?' (*Fortune*, 20 October 2020) <https://fortune.com/2020/10/20/why-do-so-few-businesses-see-financial-gains-from-using-a-i/>.

[3] According to the definition endorsed at European level by the High-Level Expert Group on Artificial Intelligence: '[a]rtificial intelligence (AI) refers to systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications).' High-Level Expert Group on Artificial Intelligence, 'A definition of AI: Main capabilities and scientific disciplines' (2019) <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341> (this link and all the following ones have been last accessed on 10 February 2024).

[4] On the concept of autonomy, see Section 2.1.

[5] For example, AI systems trained with biased data may lead to biased decisions to the detriment of minorities when screening job candidates, assessing creditworthiness for loans, or predicting criminal behavior. See J Manyika, J Silberg and B Presten, 'What Do We Do About the Biases in AI?' *Harvard Business Review* (Boston, 25 October 2019) <https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai>. Even a chatbot may turn out to be racist. This occurred with Tay, the chatbot developed by Microsoft to self-learn conversational skills and autonomously interact with users via Twitter. It was shut down in 2016. See P Lee, 'Learning from Tay's introduction' (*Official Microsoft Blog*, 25 March 2016) <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/>.

[6] Autocomplete functions of search engines may cause defamation, reputational damage, or trademark violations. See S Karapapa and M Borghi, 'Search Engine Liability for Autocomplete Suggestions: Personality, Privacy and the Power of the Algorithm' (2015) 23 International Journal of Law and Information Technology 261.

[7] For example, a flaw in the radio of a vehicle could expose the risk of unauthorised access by a third party maliciously intending to take over the control system of the self-driving car. For a similar case, see the RAPEX notification from Germany published in the EU Safety Gate website (A12/1671/15) <https://ec.europa.eu/safety-gate-alerts/screen/webReport/alertDetail/188127>. Similarly, cyberattacks on the control systems of driverless metro, autonomous weapons, industrial plants, or critical infrastructures may cause enormous damage as well, if not properly governed.

[8] On the possibility that a robo-advisor leads to wrong investments see D Litz, 'Risk, Reward, Robo-Advisers: Are Automated Investment Platforms Acting in Your Best Interest' (2018) 18 Journal of High Technology Law 367. Also, errors in automated diagnoses and surgeries may ruin a person's life. Autonomous vehicles, although promising a reduction of accidents caused by human errors, could still cause accidents due to flaws in object recognition technologies embedded in self-driving.

[9] AI can assist in diagnoses, for instance in oncology. Some AI detection systems – namely one for melanoma diagnosis – have been found to perform even better than humans. See 'Artificial intelligence will improve medical treatments' *The Economist* (London, 9 June 2018) <https://www.economist.com/science-and-technology/2018/06/09/artificial-intelligence-will-improve-medical-treatments>. Similarly, the Internet of Bodies – that is, the merger of IoT and AI with the human body – can allow patients to be reminded if they forgot to ingest their medication on the basis of a sensor implanted in their stomach. See AM Matwyshyn, 'The Internet of Bodies' (2019) 61 William & Mary Law Review 77.

[10] AI can teach itself to reduce the use of energy, required (for instance) by data centers. This is the case of the AI developed by Google's DeepMind division. On this see G Gow, 'Environmental Sustainability And AI' (*Forbes*, 21 August 2020) <https://www.forbes.com/sites/glenngow/2020/08/21/environmental-sustainability-and-ai/?sh=26938a717db3>.

[11] The Internet of Bodies can allow athletes to track their performance with wearable devices. *See* M Fierens and J De Bruyne, 'Legal and Ethical Considerations Concerning AI in Sports' (*Sports Tech Research Network*, 21 December 2020) <https://sports-tech-research-network.com/news-insights/2020/12/21/Legal-and-ethical-considerations-concerning-AI-in-sports>.

[12] For example, the Internet of Bodies can allow truck companies to check the alertness of their drivers. See Matwyshyn (n 9) 84.

[13] UNESCO, 'Recommendation on the Ethics of Artificial Intelligence' (2021) <https://unesdoc.unesco.org/ark:/48223/pf0000380455>. See also MD Dubber, F Pasquale and S Das (eds), *The Oxford Handbook of Ethics of AI* (Oxford University Press 2020); M Coeckelbergh, *AI Ethics* (The MIT Press 2020). With specific regard to algorithmic transparency and the need to shift from a black-box society to an intelligent one, see F Pasquale, *The Black Box Society* (Harvard University Press 2015). ('Rather than contort ourselves to fit 'an impersonal economy lacking a truly human purpose,' we might ask how institutions could be re- shaped to meet higher ends than shareholder value […]. Black box services are often wondrous to behold, but our black box society has become dangerously unstable, unfair, and unproductive. Neither New York quants nor California engineers can deliver a sound economy or a secure society. Those are the tasks of a citizenry, which can perform its job only as well as it understands the stakes').

[14] This concept, intended to emphasize the role of responsibility in shaping and promoting innovation, is gaining increasing attention among scholars and policy makers. On this, see BJ Koops and others (eds), *Responsible Innovation 2: Concepts, approaches, and Applications* (Springer 2015).

the risks generated by AI-based technologies. This need is particularly significant in those sectors, such as the financial market, where (i) the intensive use of AI-based instruments has proved itself to be able to generate the emergence of a parallel market in the supply of banking products and services[15] and (ii) the risks associated with the violation of individuals' rights (such as fair access to credit) and their implications for social welfare are extremely sensitive. Indeed, whereas in such markets AI systems are capable of radically changing operating models introducing significant opportunities in terms of efficiencies, cost reductions and customer service offerings, the risks involved in such innovations are likely to have a major impact, as any vulnerability of the financial market will expand over the economic system due to the contemporary financialization of markets.[16]

To this end, the European Union (EU) has already adopted a set of cross-sector initiatives under the AI strategy to foster the development of AI while addressing its impact on fundamental rights.[17] The Commission presented its AI package in April 2021, which included the proposal for an AI Regulation (hereafter also AI Act or Regulation) adopted in March 2024.[18] In addition, the EU has recently complemented the AI package with two further legislative instruments: a proposal for a revised Product Liability Directive (revised PLD)[19] – adopted by the Parliament on March 2024 – and a proposal for an AI Liability Directive (AILD).[20] The several measures mentioned contribute to the adoption of an AI liability regime at EU level, with the intent of establishing a framework that, by avoiding under- or over-compensation of victims, can achieve an environment of trust necessary for a development of AI that is beneficial to the economy and the society. At the same time, in the financial market, such measures shall relate and interact with sector-specific regulations in those areas where parallel regulatory proposals are in development.

In the following sections, we first discuss the main features of AI and the challenges they represent for traditional liability notions, such as damages, causation and duty of care. We then apply these concepts to the specific needs of the financial market (Section 1). We consider these notions as they are common to the majority of EU Member States, as shown in the Principles of European Tort Law developed by the European Group on Tort Law.[21] We then illustrate the development of the EU's approach to AI liability, from the established liability framework (Section 2) to the new cross-sectorial proposals and the regulatory initiatives in development in the financial and banking markets (Section 3). While an analysis of the relationship between the AI product liability framework and each regulation operating in financial market law goes beyond the scope of this research, this section will highlight how the use of AI systems represents a common technique in plurality of regulated activities within financial markets; hence, from a teleological perspective, AI liability regimes are going to operate in a close interplay with financial markets' rules. Finally, we conclude by providing an assessment of the AI liability framework and operate some considerations regarding its effective application to financial markets, and the possibility to reconcile it with the sector's rules.

This article not only navigates the intricate landscape of AI liability within the broader legal framework but also ventures beyond the state of the art by seamlessly applying the future EU AI liability regime to the dynamic realm of financial services. By scrutinizing the intersection of evolving technology and financial market regulations, we illuminate a path that extends beyond conventional boundaries, offering insights into a harmonious coexistence between AI liability principles and the nuanced intricacies of the financial sector. As we embrace the future, this research contributes a forward-thinking perspective that seeks to bridge the gap between cutting-edge AI governance and the complex tapestry of financial markets, laying the groundwork for a symbiotic relationship between innovation and regulation.

2. New technological features vis-à-vis traditional notions and regime(s) of liability

In this section, we first analyse how AI features challenge the traditional notions of liability, second examine the state of the art as to the current liability regime(s) in Europe,[22] and third focus on how the deployment of AI systems in the financial service market impacts the application of liability rules within it. We do this to set the ground for understanding the need to adopt new – and harmonized – rules and how these new rules will influence the provision of financial services.

2.1. Old dogs, new tricks – old notions, new features

The question as to whether current liability regime(s) in Europe are

[15] IHY Chiu and IG MacNeil (eds), *Research Handbook on Shadow Banking: Legal and Regulatory Aspects* (Edward Elgar 2018); S Schwarcz, 'Regulating Shadow Banking' (2012) 31 Review of Banking and Financial Law 619.

[16] T Lagoarde-Segot, 'Financialization: Towards a New Research Agenda' (2017) 51 International Review of Financial Analysis 113.

[17] For an overview of the European approach to AI, see European Commission, 'European approach to artificial intelligence' <digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>. For a first comment see R Justo-Hanani, 'The politics of Artificial Intelligence regulation and governance reform in the European Union' (2022) 55 Policy Sciences 137.

[18] Position of the European Parliament adopted at first reading on 13 March 2024 with a view to the adoption of Regulation (EU) 2024/… of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).The starting point of the process that led to the current AI Act dates back to April 2021 and encompasses the following documents: European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Fostering a European approach to Artificial Intelligence' COM (2021) 205 final (Communication on fostering a European approach to AI); European Commission, 'Annexes to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Fostering a European approach to Artificial Intelligence' COM (2021) 205 final - Annex (2021 Review of the Coordinated Plan on Artificial Intelligence); European Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts' COM (2021) 206 (AI Act or Regulation); European Commission, 'Commission Staff Working Document - Impact Assessment Accompanying the Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts' SWD (2021) 84 final (Impact Assessment accompanying the AI Act).

[19] Position of the European Parliament adopted at first reading on 12 March 2024 with a view to the adoption of Directive (EU) 2024/… of the European Parliament and of the Council on liability for defective products and repealing Council Directive 85/374/EEC <https://www.europarl.europa.eu/doceo/document/TA-9-2024-0132_EN.html> (Revised PLD).

[20] European Commission, 'Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)' COM (2022) 496 final (AILD).

[21] See Francesco D. Busnelli and others, *Principles of European Tort Law: Text and Commentary* (Springer 2005); K Oliphant, 'Cultures of Tort Law in Europe' (2012) 3 Journal of European Tort Law 147; M Bussani and M Infantino, 'Tort Law and Legal Cultures' (2015) 63 American Journal of Comparative Law 77.

[22] We use the plural 'regimes' as this is not a harmonized area of law at EU level. See F Cafaggi, 'Private Regulation in European Private Law' in A Hartkamp and others (eds), *Towards a European Civil Code* (4th edn, Wolters Kluwer 2011) 91.

fit for the new digital era comes from the fact that AI presents features that are unknown to the previous generation of technologies.[23]. Indeed, the main features of AI – namely complexity, opacity, autonomy, unpredictability, openness, data-drivenness, and vulnerability – challenge the traditional liability notions of damages, causal link, and duty of care. Such features, and their impact on liability, have been widely discussed in the Report on Liability for AI and emerging digital technologies realized by the Expert Group on Liability and New Technologies.[24]

In particular, AI is, in the first place, data-driven.[25] As to the stage of training the model, issues can arise in data management and preparation, notably when the data used for the model is inaccurate, non-representative and insufficient, mirroring biases present in society, or unsecured and unprotected.[26] Inaccuracy, for instance, is often the product of inaccurate labelling, i.e. 'the process by which the training data is manually assigned class labels', as we would end up with a 'skewed ground truth' as a starting point.[27] Furthermore, to operate and self-develop, AI depends on information that is not pre-installed but generated by external or internal sources (like built-in sensors). This leads to issues whenever data is flawed or missing, due to an error in transferring the data or in relation to the source. Moreover, AI is not completed once put into circulation. It keeps developing according to subsequent inputs, such as updates and upgrades, and thus needs to interact with other systems or data sources in order to operate. Its openness "by design" permits external input via some hardware plugin or wireless connection.[28] This constant interaction with outside information is what also makes these new technologies vulnerable to cybersecurity breaches.[29]

The features of data drivenness, vulnerability and openness challenge the traditional notion of damage – such as harm to persons and properties – as they enable the harm of further categories of protected interests, such as privacy, confidential information, security and cybersecurity. Indeed, there are built-in features allowing systems to access and further interact with external information (for instance through updates). These features increase AI systems' vulnerability to cybersecurity breaches, which can cause systems to malfunction and/or modify their features in a way likely to cause harm.[30] Similarly, cybersecurity breaches could imply the leaking of data, including personal or confidential data, which is further exacerbated by systems' data-drivenness. It should be no surprise, that these characteristics are particularly relevant when AI products are used in the context of financial activities. When employed for activities such as high-frequency trading, or credit scoring, AI systems constantly acquire information from outside their internal repositories (e.g. from users' social media account[31]): in such cases, the access to noisy data is likely to lead to disadvantages for borrowers or to favor unjust discrimination, or to hinder the correct functioning of the trading algorithm.[32]

In the second place, whenever AI presents a self-learning capacity, it develops the ability to interpret the environment, interact with humans, cooperate with other actors, learn new behaviours, and execute actions without – or with limited – human intervention.[33] In this case AI becomes autonomous,[34] which in turn makes its behaviour unpredictable.[35] Indeed, many systems are designed to identify and classify new stimuli and link them to self-chosen – not pre-programmed – reactions.[36] They rely on training data and data collected while interacting with surrounding environments, which in turn alters the initial algorithms. As a result, the more external data systems can process, the more unpredictable they become. Moreover, AI systems presenting these features tend to be opaque as to their functioning due to the black box nature that they develop.[37] Opacity of AI systems may only increase in the presence

[23] In this work, the wording of the EU institutions is adopted when referring to emerging digital technologies. The category of emerging digital technologies is not fully defined and exhaustively identified in European documents on the topic, where they are indicated with the exemplificative list of 'Internet of Things (IoT), Artificial Intelligence, advanced robotics and autonomous systems', as in: European Commission, 'Commission Staff Working Document - Liability for emerging digital technologies - Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Artificial intelligence for Europe' SWD (2018) 137 final (Commission Staff Working Document - Liability for emerging digital technologies).

[24] European Commission, Directorate-General for Justice and Consumers, 'Liability for artificial intelligence and other emerging digital technologies' (Report from the Expert Group on Liability and New Technologies – New Technologies Formation, Publications Office 2019) <https://data.europa.eu/doi/10.2838/573689> (Expert Group Report on Liability for Artificial Intelligence and other Emerging Digital Technologies).

[25] N Sambasivan and others, 'Everyone wants to do the model work, not the data work: data cascades in high-stakes AI' (2021) in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Association for Computing Machinery 2021) <https://doi.org/10.1145/3411764.3445518>; S Ameen and others, 'AI Diagnostic Technologies and the Gap in Colorectal Cancer Screening Participation' in B Séroussi and others (eds), *Challenges of Trustable AI and Added-Value on Health - Proceedings of MIE 2022* (Studies in Health Technology and Informatics vol 294, IOS Press 2022) 803-804, maintaining that '[d]ata is the critical infrastructure necessary to build AI systems', with issues arising from 'AI/ML practices that undervalue data quality.'

[26] M Jacobs and J Simon, 'Assigning Obligations in AI Regulation: A Discussion of Two Frameworks Proposed by the European Commission' (2022) 1(1) Digital Society <https://doi.org/10.1007/s44206-022-00009-z>.

[27] S Barocas and AD Selbst 'Big data's disparate impact' (2016) 104 California Law Review 671.

[28] These open systems come as hybrid combinations of hardware, software, ongoing updates and services, as explained in M Geistfeld and others (eds), *Civil Liability for Artificial Intelligence and Software* (De Gruyter 2023) 549-550.

[29] A Lohn, 'Hacking AI: A Primer for Policymakers on Machine Learning Cybersecurity' (Center for Security and Emerging Technology 2020) <https://doi.org/10.51593/2020CA006>.

[30] Geistfeld and others (n 28) 551.

[31] N Packin and Y Lev-Aretz, 'On Social Credit and the Right To Be Unnetworked' (2016) Columbia Business Law Review 339.

[32] D Ruck, 'Tricking the Trade. How High Frequency Trading AI Can Be Manipulated by Adversarial AI' (*Medium*, 7 November 2020) <https://medium.com/swlh/tricking-the-trade-4faadd211113>.

[33] When AI is based on supervised machine learning, it is trained on labelled datasets. Instead, when AI is based on unsupervised ML, it learns from unlabeled data, through a pattern-seeking approach. As for reinforcement learning, it works towards a goal through trial and error. See CS Smith, 'Computers already learn from us. But can they teach themselves?' *The New York Times* (New York, 8 April 2020) <https://www.nytimes.com/2020/04/08/technology/ai-computers-learning-supervised-unsupervised.html>. Examples of self-learning AI models can be found, for instance, in the medical field: R Matheson, 'Artificial intelligence model 'learns' from patient data to make cancer treatment less toxic' (*MIT News*, 9 August 2018) <https://news.mit.edu/2018/artificial-intelligence-model-learns-patient-data-cancer-treatment-less-toxic-0810>.

[34] For the development of autonomy in the medical sector see D Bitterman, H Aerts and R Mak, 'Approaching autonomy in medical artificial intelligence' (2020) 2 The Lancet Digital Health 447.

[35] Autonomous capabilities and intelligence ungoverned by human directions or supervision could lead to unexpected outcomes, as shown by the story of Alice and Bob, i.e., two chatbots developed to learn autonomous bargaining skills that started to interact using their own code, indecipherable for humans. See T Simonite, 'No, Facebook's Chatbots Will Not Take Over the World' *Wired* (San Francisco, 1 August 2017) <https://www.wired.com/story/facebooks-chatbots-will-not-take-over-the-world/>.

[36] This is the case of unsupervised learning algorithms. Consider, for example, an AI system trained to detect anomalies, where the system would learn to identify them without prior knowledge of what constitutes normal behavior. Such systems range from fraud detection to medical diagnosis.

[37] Pasquale (n 13). More recently, B Vaassen, 'AI, Opacity, and Personal Autonomy' (2022) 35 Philosophy & Technology 88.

of self-learning features, as algorithms no longer come as readable code but amount to black boxes that are almost impossible to understand. In addition, AI systems can also present a high degree of complexity whenever there is interdependency between different components and layers.[38] This increases the variety of players involved and complicates the understanding of potentially harmful processes. Whereas cybersecurity and privacy-related problems are common to every implementation of an AI system – and, more generally, many of the issues associated with financial institutions' use of AI are quite similar to those posed in other areas[39] – the issue of AI autonomy and opacity is of particular significance in financial markets due to the specific characteristics of the harm that biased or unfair processes might cause. It has been observed that growing usage of AI by financial institutions has major implications for financial stability: inter alia, unstructured and semi-structured data sources in processes such as the analysis of credit scoring applications could produce unjust discriminations in access to credit,[40] determine phenomena of disparate impact and financial exclusion towards minorities[41] and, ultimately, misallocate financial resources and undermine the role of capital markets as redistributive mechanisms.[42] Similarly, using AI to devise trading and investment strategies for clients' portfolio management can lead – when improperly operated – to tight coupling, intensification of volatility, ripple effects and, more in general, amplify systemic risks and lead to flash crashes.[43]

Besides the deployment of AI systems in B2C interactions, risks are also present and significant when considering that these technologies are implemented in model risk management and structural market analyses as well[44]: as AI is used to evaluate how well banks' risk models are performing, and how they should develop their business in the future, the structural implications of a potential malfunctioning are critical.[45]

The features of autonomy, unpredictability, opacity and complexity challenge the notions of causation and duty of care.[46] As to the former, liability regimes pivot around the principle that the victim should prove that the damage originated by some conduct or risk is attributable to the defendant. However, providing evidence of causation can become difficult with interconnected devices like automated vehicles (AVs) (combining hardware, software, connectivity and data) or self-learning AI systems (fuelled by machine learning – including deep learning – techniques and based on multiple external data collection).[47] AI-empowered products may act in ways that were not envisaged at the time that the system was first put into operation, and these behaviours may be so autonomous to make difficult for a victim to prove that their harm is attributable to a tortfeasor (and thus to find compensation for their harm) as the AI's unpredictable conduct dilutes the causal link.[48] The combination of opacity and complexity leads to issues in establishing causation, due to the variety of actors and of causes (possibly successive) which could have contributed to the damage.[49]

As to the latter, a duty of care can be deemed central in most Member States' fault-based liability regimes and requires the adherence to a standard of reasonable care while performing any acts that could foreseeably harm others.[50] While statutory language may in certain cases define such duties,[51] in many others they are reconstructed by the court based on social beliefs about the prudent and reasonable course of action in the circumstances at stake.[52] Applying fault-based liability rules to AI

[38] These components can range from tangible parts and devices (e.g. sensors, actuators, hardware), to software components, data, and connectivity features. This is described, for EDTs in general, in the Commission Staff Working Document - Liability for emerging digital technologies (n 23).

[39] J Prenio and J Yong, 'Humans Keeping AI in Check – Emerging regulatory expectations in the financial sector' (FSI Insights on policy implementation no. 35, Financial Stability Institute 2021)<https://www.bis.org/fsi/publ/insights35.pdf>.

[40] A Davola, 'Technological innovation in creditworthiness assessment' (2019) 5(2) Open Review of Management, Banking and Finance, <https://openreviewmbf.org/2019/10/10/technological-innovation-in-creditworthiness-assessment/>.

[41] N Geslevich-Packin and Y Lev-Aretz, 'On Social Credit and the Right to Be Unnetworked' (2016) Columbia Business Law Review 339; N Campisi and C Lupini, 'From Inherent Racial Bias to Incorrect Data —The Problems With Current Credit Scoring Models' (*Forbes Advisor*, 26 February 2021) <https://www.forbes.com/advisor/credit-cards/from-inherent-racial-bias-to-incorrect-data-the-problems-with-current-credit-scoring-models/>.

[42] See in general G Epstein, 'Introduction: Financialization and the World Economy' in G Epstein (ed), *Financialization and the World Economy* (Edward Elgar Publishing 2005).

[43] D Sornette and S von der Becke, 'Crashes and high frequency trading' (2011) Swiss Finance Institute Research Paper No. 11–63; R Feldman and K Stein, 'AI Governance in the Financial Industry' (2022) 27 Stanford Journal of Law, Business, and Finance.

[44] OECD, 'Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers' (2021) <https://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf>.

[45] D Arner and others, 'Fintech, Regtech and Systemic Risk: The Rise of Global Technology Risk' in D Arner, E Avgouleas, D Busch, and S Schwarcz (eds), *Systemic Risk in the Financial Sector: Ten Years after the Great Crash* (McGill-Queen's University Press 2019).

[46] Geistfeld and others (n 28).

[47] Expert Group Report on Liability for Artificial Intelligence and other Emerging Digital Technologies (n 24).

[48] For example, in the Expert Group Report on Liability for Artificial Intelligence and other Emerging Digital Technologies (n 24) 20, the group of experts provides the example of a smoke detector that fails to activate the alarm. If this is due to a problem with the wiring, it is easy to provide evidence, but if the failure is linked to a firmware error, proving it may not be as straightforward. Even if there is evidence that the alarm did not go off, proving that it is because of a firmware error requires a careful examination of the code and its compatibility with the smoke detector's hardware. The challenge intensifies even more when the smoke detector's algorithm is created or altered by an AI system using machine learning and deep learning techniques. This AI system learns from various external data since its inception, therefore, even if there are no changes to the original software design, understanding the embedded criteria guiding data collection, analysis, and decision-making is not easily explainable and often demands costly expert analysis. This practical hurdle becomes a significant obstacle for someone seeking compensation for the unpredictability of the chances of success poses a challenge for the victim upfront. This example shows that 'the autonomy and self-learning capacity of the technology may be seen as breaking the causal link between the actor's conduct and the damage', which is the problem 'the problem of attribution of the operation and its outcome to a person, which should be solved by legally ascribing all the emerging digital technology's actions and their effects to the operator of the technology' (Expert Group Report on Liability for Artificial Intelligence and other Emerging Digital Technologies (n 24) 54).

[49] BA Koch and others, 'Response of the European Law Institute to the Public Consultation on Civil Liability – Adapting Liability Rules to the Digital Age and Artificial Intelligence' (2022) 13 Journal of European Tort Law 25.

[50] Geistfeld and others (n 28). As to the scope of a duty of care across jurisdictions, see E Büyüksagis and F Werro, 'The Bounds between Negligence and Strict Liability' in M Bussani and AJ Sebok (eds), *Comparative Tort Law. Global Perspectives* (Edward Elgar 2021) 204-205.

[51] For example, as pointed out by E Büyüksagis and F Werro (n 50) 204, German law establishes a duty of care specifically for legally protected interests, including life, property, and personality, as outlined in § 823 of the Civil Code (BGB). Unless there is harm to one of these fundamental interests or a breach of a particular protective rule (Schutznorm), the German tort system generally does not allow compensation for 'pure' economic losses (reine Vermögensschäden). In contrast, French law does not make such distinctions and considers a general duty of care for all situations where a lack of caution could foreseeably lead to harm (Article 1240 of the French Civil Code).

[52] While a complete reference to all the cases that over the years have contributed to shape the duty of care within the Member States is not possible, many examples can be found in Helmut Koziol's books: H Koziol, *Basic Questions of Tort Law from a Germanic Perspective* (Sramek Verlag 2012); H Koziol, *Basic Questions of Tort Law from a Comparative Perspective* (Sramek Verlag 2015).

systems is difficult, because they lack well-established models of proper functioning and develop by learning without direct human control. The processes running them cannot all be measured according to duties of care designed for human conduct, an accepted standard of care for the creation and operation of autonomous systems has not emerged yet. It may sometimes be hard even to identify the person obliged to meet such duty of care. In fact, it could be arbitrary to assign liability for any damage caused by an AI product to the designer of the algorithm. Depending on circumstances liability should be allocated also to other subjects or entities, such as operators,[53] yet this is not self-evident in the case of AI systems.[54] However, according to any liability regime, tracing a damage back to a specific person is still a fundamental prerequisite for any fault-based claim.[55]

All the above-mentioned features make it more difficult to assess liability and compliance with applicable legislation unless they are adequately governed.[56] A first answer to the challenges above-illustrated comes from the scholars that suggest to update the traditional liability notions to align them to the technological pace.[57] In particular, applying common enterprise liability to cases involving AI systems would imply joint and several liability of all subjects involved in the design, programming and deployment of an AI application.[58] Although this could facilitate claims of compensation for damage; it might be ineffective in allocating costs and setting prevention incentives for all relevant players.[59]

Other scholars, instead, urge to reconceptualize intelligent and autonomous machines as entities with the status of a "person" under the law, such that AI can be held directly liable for harm – just as legal entities are.[60] They argue that an intelligence "even able to supersede humans in a number of areas" could sometimes be at fault. This legal fiction, however, may open up more problems than it solves, particularly as to the definition of selection criteria and equity requirements, as well as to the allocation of costs among all parties involved in the development and use of AI applications.[61] So far, legislators seem far from revolutionizing the traditional notions of liability to introduce some sort of robot's fault.[62]

There have been also proposals to apply a "reasonable algorithm" standard to self-learning systems, given their similarity to humans in decision-making and the consequent damage. This solution too poses a crucial, so far unresolved, question: what could be considered reasonable behavior for algorithms?[63] Rather than resorting to conceptually new theories, another – maybe more viable – option that has been proposed is that of introducing a predetermined, detailed and acceptable level of care (or quasi-safe-harbor) for designers, manufacturers, owners and users of AI.[64] If the level of care is unmet, a presumption of negligence and, therefore, liability would be triggered; if met, the defendant would enjoy a quasi-safe harbor, while the claimant would bear the burden of proving actual negligence.[65]

2.2. Traditional liability regime(s)

The notions of damage, causality and duty of care above illustrated are rooted in the national liability regimes of Member States as at EU level a liability framework is only partially harmonised.[66] Indeed, the existing EU tort law rules are currently limited – at least until the entry into force of the framework on AI liability[67] – to the current version of product liability under Directive 85/374/EC,[68] liability for infringing data protection law (under Article 82 of the GDPR),[69] and liability for infringing competition law (under Directive 2014/104/EU).[70]

[53] In this paper, operator is defined as per the AI Act (n 18), art 3(8): 'operator' means the 'provider, product manufacturer, deployer, authorised representative, importer or distributor'.

[54] For instance, with regard to autonomous weapons, 'somehow human responsibility and accountability for the actions taken by the machine evaporate and disappear. The soldier in the field cannot be expected to understand in any serious way the programming of the machine; the designers and programmers operate on a completely different legal standard; the operational planners could not know exactly how the machine would perform in the fog of war; and finally, there might be no human actors left standing to hold accountable' K Anderson and MC Waxman, 'Debating Autonomous Weapon Systems, their Ethics, and their Regulation under International Law' in R Brownsword, E Scotford and K Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology* (Oxford University Press 2017).

[55] European Commission, 'White Paper On Artificial Intelligence - A European approach to excellence and trust' COM (2020) 65 final (White Paper On Artificial Intelligence).

[56] In the financial sector, this is clearly acknowledged in ESMA (European Securities and Markets Authority), 'Artificial Intelligence in EU securities markets' (2023) <https://www.esma.europa.eu/sites/default/files/library/ESMA50-164-6247-AI_in_securities_markets.pdf>.

[57] See Jacobs and Simon (n 26); Geistfeld and others (n 28); BA Koch and others (n 49).

[58] DC Vladeck, 'Machines Without Principals: Liability Rules and Artificial Intelligence' (2014) 89 Washington Law Review 117; B Chan, 'Applying a common enterprise theory of liability to clinical AI systems' (2021) 47 American Journal of Law and Medicine 351.

[59] G Comandè, 'Multilayered (Accountable) Liability for Artificial Intelligence' in S Lohsse, R Schulze, and D Staudenmayer (eds), *Liability for Artificial Intelligence and the Internet of Things* (Hart-Nomos 2019) 165.

[60] See J Hage, 'Theoretical foundations for the responsibility of autonomous agents' (2017) 25 Artificial Intelligence and Law 255; BW Jackson, 'Artificial Intelligence and the Fog of Innovation: A Deep-Dive on Governance and the Liability of Autonomous Systems' (2019) 35 Santa Clara High Technology Law Journal 35. This debate has a long history, as shown by LB Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 North Carolina Law Review 1231. A case against treating robots like humans is made by H Eidenmüller, 'The Rise of Robots and the Law of Humans' (2017) Oxford Legal Studies Research Paper No.27/2017.

[61] G Comandè, 'Intelligenza artificiale e responsabilità: Il carattere trasformativo dell'IA e il problema della responsabilità' [2019] Analisi giuridica dell'economia 169.

[62] In the Expert Group Report on Liability for Artificial Intelligence and other Emerging Digital Technologies (n 24), the Expert Group denied the necessity to adopt the notion of electronic personhood.

[63] KA Chagal-Feferkorn, 'How Can I Tell if My Algorithm Was Reasonable?' (2021) 27 Michigan Technology Law Review 213 <https://repository.law.umich.edu/mtlr/vol27/iss2/2>. See JS Borghetti, 'Civil Liability for Artificial Intelligence: What Should its Basis Be?' (2019) 17 La Revue des Juristes de Sciences Po 94 <https://ssrn.com/abstract=3541597> for a discussion on the 'test of the reasonable robot' and on the shortcomings of comparing an algorithm's output with a reasonable human behaviour.

[64] See O Rachum-Twaig, 'Whose Robot Is It Anyway?: Liability for Artificial-Intelligence-Based Robots' (2020) 11 University of Illinois Law Review 1141, in particular 1172-73.

[65] ibid.

[66] Oliphant (n 21); Bussani and Infantino (n 21).

[67] See Section 3.

[68] Council Directive 85/374/EEC of 25 July 1985 on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability for Defective Products [1985] OJ L210/29 (PLD). A thorough account of such instrument is made in D Fairgrieve and others, 'Product Liability Directive' in P Machnikowski (ed), *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies* (Intersentia 2016) 17. See also MG Faure, 'Economic Analysis of Product Liability' in P Machnikowski (ed), *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies* (Intersentia 2016) 619.

[69] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1 (GDPR).

[70] Directive 2014/104/EU of the European Parliament and of the Council of 26 November 2014 on Certain Rules Governing Actions for Damages under National Law for Infringements of the Competition Law Provisions of the Member States and of the European Union [2014] OJ L349/1.

As a result, harmful damages arising during the use of AI systems are likely to be compensated under existing national tort and contract law, or, where applicable, through specific liability provisions of Member States. For example, some national jurisdictions have specifically regulated the use of AVs, also providing for coverage of any damages caused, by insurance or by reference to the general rules.[71]

In general, domestic tort laws include a rule introducing fault-based liability with a broad scope of application, accompanied by several more specific rules which either modify the premises of fault-liability (especially in the distribution of the burden of proof) or establish liability independently from fault (strict or risk-based liability).[72] Most liability regimes also encompass the notion of liability for others (indirect or vicarious liability), which can, in turn, be fault- or risk-based, depending on the case or the country.[73] While this is not the place to engage in a comparative analysis of each Member States liability framework, it is worth pointing out that they all share some common principles.[74] A general rule of fault-based liability is a common ground.[75] When an actor fails to take due care, and this negligence causes harm to another – or they cause such harm intentionally – this actor is liable to compensate the victim. Usually, what triggers liability is harm to the fundamental interests of a person, such as life, health, bodily integrity, freedom of movement, private property, and in some countries also purely economic losses and harm to human dignity.[76]

In addition, all Member States' legal systems encompass product liability as a result of the current PLD implementation which however dates back to 1984 - which will be replaced by the revised PLD once finally approved. On this base, a damage claim for harm generated by a defective product does not require a finding of fault on the part of the manufacturer, as, in principle, the PLD should introduce a harmonized strict – not fault-based – liability regime for defective products.[77] However, the regime that the PLD introduces resembles more a watered-down version of negligence liability than a strict liability regime since a claimant must, in any case, prove the defect and that such defect generates the harm that she suffered.[78] In sum, the current product liability regime only covers damages generated by defective products, leaving outside the provision of services, for which then the default negligence-based regime revives.[79] Moreover, the PLD implementation not only has not been consistent in all Member States, but it lacks to cover instances generated by the use of AI,[80] and, more in general, all the so called emerging digital technologies (EDTs).[81]

As a result, the EU scenario in force before the last wave of proposals is quite fragmented. Disparities in Member States' legislation and case-law concerning liability (fault-based, strict and vicarious) may produce distortions of competition and impair the functioning of the single digital market, while the moderate pace of European legislative harmonization may no longer be suitable to the rapid changes brought by AI.[82]

2.3. The relevance of liability rules in the financial service market

Liability regimes have traditionally played a major role in financial markets as well, being widely used by the EU legislator as a substantive regulatory instrument. Prominent examples include, inter alia, the regulation of the liability of credit rating agencies,[83] the breach of regulatory duties identified by PSD2[84] and MiFID II[85] respectively, and – more recently – the rules on civil liability for crypto-asset issuers under the new Markets in Crypto-Assets Regulation.[86] Within financial markets, liability rules operate both as a compensatory device and as a deterrent against violations of the standards of conduct set by the legislator for financial market agents and entities. Regarding this aspect, it is worth pointing out that even if the deterrent function of tort law is not unequivocally accepted within the territory of the European Union, it seems unquestionable that in a context such as the financial market, which is characterized by a significant focus on prudential supervision, the ability of a regulation to prevent ex ante the harmful effect assumes particular relevance. This, even considering the systemic effect that an AI malfunction could have on financial markets. Hence, it should come with no surprise, that the interplay between financial regulation and civil liability in European law is traditionally acknowledged to be pivotal in the EU multi-level system of governance.[87]

Therefore, in consideration of such complementary function that liability plays, it is no surprise that the abovementioned problems deriving from the deployment of AI systems have a major impact on

[71] See for example the Germany amended its Road Traffic Act to allow driverless vehicles on public roads: Act Amending the Road Traffic Act and the Compulsory Insurance Act (Autonomous Driving Act), July 2021. Bundesgesetzblatt Jahrgang 2021 Teil I Nr. 48, ausgegeben zu Bonn am 27. Juli 2021 <https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl121s3108.pdf#_bgbl_%2F%2F*%5B%40attr_id%3D%27bgbl121s3108.pdf%27%5D_1676679941289>. In France, too, the framework for the deployment of automated vehicles was initiated through a decree: Décret n° 2021-873 du 29 juin 2021 portant application de l'ordonnance n° 2021-443 du 14 avril 2021 relative au régime de responsabilité pénale applicable en cas de circulation d'un véhicule à délégation de conduite et à ses conditions d'utilisation <https ://www.legifrance.gouv.fr/jorf/id/JORFTEXT000043729532>. See, among many, R Diehl and MI Thue, 'Autonomous Vehicle Testing Legislation: A Review of Best Practices from States on the Cutting Edge' (2017) 21 Journal of Technology Law & Policy 197; and, in the US: MA Geistfeld, 'A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation' (2017) 105 California Law Review 1611.

[72] Bussani and Sebok (n 50).

[73] ibid.

[74] Oliphant (n 21); Bussani and Infantino (n 21).

[75] European Group on Tort Law, 'Principles of European Tort Law' art 1:101. For a comment, see FD Busnelli and others, *Principles of European Tort Law: Text and Commentary* (Springer 2005).

[76] M Infantino, 'Protected Interests under the Principles of European Tort Law (Art 2:102 PETL) – Preserving the Past for Shaping the Future' (2023) 14 Journal of European Tort Law 42.

[77] PLD (n 68) recital 2 'liability without fault'.

[78] PLD (n 68) art 4.

[79] PLD (n 68) art 1.

[80] For a survey of the issues as to the application of the Product Liability Directive to EDTs, see C De Meeus, 'The Product Liability Directive at the Age of the Digital Industrial Revolution: Fit for Innovation?' (2019) 8 Journal of European Consumer and Market Law 149.

[81] For instances of the gaps posed by the PLD in terms of EDTs, see BA Koch and others (n 49); S Lohsse and others (n 59).

[82] J Morais Carvalho and K Nemeth, 'Time for a Change? Product Liability in the Digital Era' (2019) 8 Journal of European Consumer and Market Law 160.

[83] See B Haar, 'Civil Liability of Credit Rating Agencies after CRA 3 – Regulatory All-or-Nothing Approaches between Immunity and Over-Deterrence' (2013) University of Oslo Faculty of Law Research Paper No. 2013-02 <https://ssrn.com/abstract=2198293>; G Deipenbrock, 'The European Civil Liability Regime for Credit Rating Agencies from the Perspective of Private International Law – Opening Pandora's Box?' (2015) Special Issue of International and Comparative Corporate Law Journal, on Civil Liability of Credit Rating Agencies in the European Union – Selected Legal and Economic Aspects, edited by Gudula Deipenbrock and Mads Andenas, University of Oslo Faculty of Law Research Paper No. 2015-02<https://ssrn.com/abstract=2546268>.

[84] Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market, amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC [2015] OJ L337/35.

[85] Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU [2014] OJ L173/349.

[86] See Section 3.3.

[87] O Cherednychenko, 'Financial regulation and civil liability in European law' in O Cherednychenko & M Andenas (eds), *Financial Regulation and Civil Liability in European Law* (Edward Elgar 2020).

financial markets as well. The growing interest toward the effect that liability rules are likely to have on financial markets stems as a necessary by-product of the growing usage of tech-based solutions for the provision of financial services: as financial service providers leverage digital technologies to revolutionize their financial product offerings and the markets in general, AI has emerged as pivotal, for instance, in credit scoring-related activities, with decision-making models exploiting AI to analyse large amounts of alternative data to determine a borrower's credit risk.[88] Yet – and even with credit scoring being probably the most analysed example in legal scholarship – applications of AI in financial services operate across a much vaster spectrum, ranging from High-Frequency Trading,[89] to the organization of compliance and risk management,[90] to the usage of chatbots and (more in general) of robo-advisory services for clients.[91]

The more AI applications become used, the more stringent becomes the connection between financial markets regulation and liability rules and, henceforth, the need for a more in-depth analysis of the relationship between the two. As finance becomes increasingly tech-based, potential malfunctioning or misuses of the technologies underlying the provision of products and services are likely to have a "domino effect" on interactions occurring on the market. While this is a common problem occurring in a plurality of markets, the concern is particularly significant in financial markets due, on the one hand, to the heavy reliance of this sector on ICTs and, more in general, digital technologies and, on the other hand, to the relevance of financial transactions as a backbone for the development of contemporary economies.[92] It should therefore not surprise that, in parallel with the development of the debate on the regulation of AI liability, an analogous conversation has been occurring on how different regulatory frameworks could affect the provision of financial products and services.[93]

3. The complementarity between AI governance and the AI liability regime

The recent efforts to address the AI liability issue at EU level have resulted in the adoption of several legislative instruments that, combined, give rise to a European AI liability framework. Particularly, these instruments are the AI Act,[94] the revised PLD[95] and the proposal for a brand-new AILD,[96] the former as a means of ex ante regulation, the latter two as means of ex post regulation. In fact, while the AI Act introduces provisions to govern AI, particularly high-risk AI systems, the directives propose rules to face the scenario in which a lack of AI governance generated damages.

In this section, we first illustrate the evolution of the EU policies as they highlight the importance for the regulatory framework to account for the legal issues raised by AI (including questions of liability) and have led to the adoption of hard law. Second, we analyse both the AI Act and the directives to analyse their complementarity and evaluate if they really achieve the objective of contributing to the creation of an environment of trust. Finally, we address the impact of this regulatory framework in the making on the financial markets to underscore, on one hand, the practical overlap between these two phenomena, and to emphasize, on the other hand, the crucial need for effective coordination between these regulations to advance investor protection successfully.

3.1. The role of EU policies in shaping the EU AI liability regime

The debate on whether the current liability regime is fit for accommodating the issues previously described has been quite lively within the EU, in particular as to what extent the existing liability schemes are adapted to the emerging market realities that follow the development of EDTs in general, and AI in particular. At policy level the issue has been tackled in two waves of subsequent policy documents. A first wave introducing the need to achieve an environment of trust for EDTs to flourish, and a second more focused on the establishment of a framework of excellence and trust for AI in particular, analysed in the following.

3.1.1. Towards an environment of trust for AI

In February 2017, the Resolution on Civil Law Rules on Robotics with recommendation to the Commission[97] proposed a whole range of legislative and non-legislative initiatives in the field of robotics and AI. A year later, in February 2018, the European Parliamentary Research Service study on a common EU approach to liability rules and insurance for connected and autonomous vehicles[98] was adopted as an added value assessment accompanying the Resolution on Civil Law Rules. On April 25 2018, the Commission published a Staff Working Document on Liability for Emerging Digital Technologies,[99] accompanying the Commission's Communication on Artificial Intelligence for Europe,[100] which provides the starting point of the discussions on liability in the context initially of EDTs in general and later specifically on AI.

All these documents, as well as the following Sibiu Communication of May 2019,[101] stress that a robust regulatory framework should address the ethical and legal questions surrounding AI, including those related to liability. In its 2018 Communication on AI, the Commission also announced the adoption of a report assessing the implications of EDTs on existing

[88] Davola (n 40); N Aggarwal, 'The norms of algorithmic credit scoring' (2021) 80 The Cambridge Law Journal 42; ML Montagnani and C Paulesu, 'Towards an Ecosystem for Consumer Protection in the Context of AI-based Credit Scoring' (2022) 33 European Business Law Review 557.

[89] D Busch, 'MiFID II: regulating high frequency trading, other forms of algorithmic trading and direct electronic market access' (2016) 10 Law and Financial Markets Review 72; F Consulich, M Maugeri, C Milia and others, 'AI and market abuse: do the laws of robotics apply to financial trading?' (2023) CONSOB Legal Research Papers no. 29.

[90] E Kurshan, H Shen and J Chen, 'Towards self-regulating AI: challenges and opportunities of AI model governance in financial services' (2021) Proceedings of the First ACM International Conference on AI in Finance (ICAIF '20) 1 <https://doi.org/10.1145/3383455.3422564>; N Remolina, 'Generative AI in Finance: Risks and Potential Solutions' (2023) Singapore Management University School of Law Research Paper Forthcoming, SMU Centre for AI & Data Governance Research Paper Forthcoming <https://ssrn.com/abstract=4628235>.

[91] C Hammer, 'Regulation of Robo-Advisory in Europe and Germany' (2021) in P Scholz (ed) *Robo-Advisory* (Palgrave Macmillan 2020); R Buckley and others, 'Regulating Artificial Intelligence in Finance: Putting the Human in the Loop' (2021) 43 Sydney Law Review 43.

[92] W Currie and T Lagoarde-Segot, 'Financialization and Information Technology: Themes, issues and Critical Debates – part I' (2017) 32 Journal of Information Technology 211.

[93] T Lin, 'Artificial Intelligence, Finance, and the Law' (2019) 88 Fordham Law Review 531.

[94] AI Act (n 18).

[95] Revised PLD (n 19).

[96] AILD (n 20).

[97] European Parliament Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) [2018] OJ C252/239.

[98] T Evas, 'A common EU approach to liability rules and insurance for connected and autonomous vehicles. European Added Value Assessment Accompanying the European Parliament's Legislative Own-initiative Report' (European Parliament - European Parliamentary Research Service 2018) 615-635.

[99] Commission Staff Working Document - Liability for emerging digital technologies (n 23).

[100] European Commission, 'Artificial intelligence for Europe' (Communication) COM (2018) 237 final.

[101] European Commission, 'Europe in May 2019: Preparing for a more united, stronger and more democratic Union in an increasingly uncertain world - The European Commission's contribution to the informal EU27 leaders' meeting in Sibiu (Romania) on 9 May 2019' (Communication) COM (2019) 218 final.

safety and liability frameworks by mid-2019. In its 2019 Work Programme, it confirmed it would "continue work on the emerging challenge of Artificial Intelligence by enabling coordinated action across the European Union."[102] Accordingly, on April 2019, the high-level Expert Group on Artificial Intelligence set up by the European Commission listed liability frameworks among the non-technical methods for securing and maintaining a lawful and trustworthy AI,[103] on the assumption that an environment of trust is crucial for fully reaping the benefits of innovation.[104]

In order to provide an answer on how the liability regime could assist in achieving an environment of trust, in March 2018, the Commission also set up an Expert Group on Liability and New Technologies, operating in two different formations: the Product Liability Directive formation and the New Technologies formation.[105] This second formation was in particular asked to assess "whether and to what extent existing liability schemes are adapted to the emerging market realities following the development of the new technologies such as Artificial Intelligence, advanced robotics, the IoT and cybersecurity issues."[106] The experts were requested to examine whether the current liability regimes are still "adequate to facilitate the uptake of […] new technologies by fostering investment stability and users' trust."[107] In case of shortcomings, the expert group was invited to make recommendations for amendments, without being limited to existing national and EU legal instruments. However, recommendations were to be limited to matters of extracontractual liability, leaving aside in particular corresponding (and complementary) rules on safety and other technical standards. As a result of the expert group's activity, in November 2019 the Report "Liability for Artificial Intelligence and other Emerging Digital Technologies" was published.[108] This undertook an assessment of existing liability regimes in the wake of emerging technologies and it concluded that the current ones in force in the Member States ensured at least basic protection of victims whose damage is caused by the operation of such new technologies, while also hinting to some adjustments that might be needed.

3.1.2. A European approach to artificial intelligence

The need for some adjustments was confirmed also in the White Paper on artificial intelligence to foster excellence and trust[109] and in the Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics.[110] Both documents stress that the ultimate goal is to ensure remediation of damage caused by AI and overall reliability, while promoting investment stability and, more generally, innovation. In this context, efficient liability rules are deemed paramount for trustworthiness, which in turn is a prerequisite for the uptake of AI. Pursuing such a strategy was also defined a crucial step to strengthen European technology sovereignty and affirms the role of the EU on the international stage as "the most attractive, secure and dynamic data-agile economy in the world."[111]

To achieve these goals, the European Commission suggested a regulatory and investment-oriented approach, entailing, among other things, adjustments to current European and national liability regimes. Indeed, a fragmented legal landscape sprinkled of different national initiatives could lead to the fragmentation of the single market and, consequently, endanger not just legal certainty, but also the emergence of a dynamic and flourishing European industry. Hence, the European Commission stressed the importance of aligning the efforts at European, national, and regional level,[112] while promoting partnership between the private and the public sector towards an "ecosystem of excellence" with proper incentives to research, innovation and deployment,[113] an "ecosystem of trust" duly protecting fundamental rights and consumers' rights[114] such as privacy

[102] European Commission, 'Commission Work Programme 2019: delivering what we promised and preparing for the future' (Communication) COM (2018) 800 final.

[103] European Commission, Directorate-General for Communications Networks, Content and Technology, 'Ethics guidelines for trustworthy AI' (High-Level Expert Group on Artificial Intelligence, Publications Office 2019) <https://data.europa.eu/doi/10.2759/346720>.

[104] European Commission, 'Building Trust in Human-Centric Artificial Intelligence' (Communication) COM (2019) 168 final.

[105] See European Commission, 'Expert Group on liability and new technologies (E03592)' (Register of Commission Expert Groups and Other Similar Entities) <https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&groupId=3592&fromMeetings=true&meetingId=7910>.

[106] See European Commission, 'Call for Applications for the Selection of Members of the Expert Group on Liability and New Technologies (E03592)' <https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&groupId=3592&fromMeetings=true&meetingId=7910>.

[107] ibid.

[108] Expert Group Report on Liability for Artificial Intelligence and other Emerging Digital Technologies (n 24).

[109] White Paper on Artificial Intelligence (n 55).

[110] European Commission, 'Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics' COM (2020) 64 final.

[111] White Paper on Artificial Intelligence (n 55) 3.

[112] Stronger coordination is encouraged in European Commission, 'Coordinated Plan on Artificial Intelligence' (Communication) COM (2018) 795 final, regarding a plan among the European Commission, Member States, Norway, and Switzerland for some 70 joint actions in the following key areas: (i) increasing investment, (ii) making more data available, (iii) fostering talent, and (iv) ensuring trust. The plan will run until 2027, with regular monitoring and update.

[113] To foster investments, the European Commission has proposed a number of measures under the Digital Europe Programme, Horizon Europe and the Multiannual Financial Framework for 2021 to 2027. On this, see European Commission, 'Info session Horizon 2020: Artificial intelligence for manufacturing' (18 November 2019) <https://ec.europa.eu/digital-single-market/en/news/info-session-horizon-2020-artificial-intelligence-manufacturing>. A key role is recognized to Digital Innovation Hubs, see European Commission, 'Digital Innovation Hubs: helping companies across the economy make the most of digital opportunities' (12 January 2021) <https://digital-strategy.ec.europa.eu/en/library/digital-innovation-hubs-helping-companies-across-economy-make-most-digital-opportunities-brochure>. According to the European Commission, making the European Union a lighthouse center of research requires also upskilling the workforce, offering world-leading masters programs, and attracting the best professors and scientists. See White Paper on Artificial Intelligence (n 55) 7.

[114] Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive') [2005] OJ L149/22; Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council [2011] OJ L304/64.

and non-discrimination,[115] and through liability rules.

In line with the Report from the expert group, the European Commission's analysis of the current legal frameworks concluded for the adaptations of current norms and the adoption of new specific legislation, pursuing a targeted, risk-based approach, and ensuring effective enforcement. In order to address both current and anticipated technological, societal and commercial developments, such revised regulatory framework aims at effectively balance protection and innovation, while not being excessively prescriptive and burdensome for businesses.

3.2. A European AI liability framework in the making

To fill the gap generated by AI specificity in the EU's existing liability regime, the EU institutions have been working on a regulatory framework fit for AI liability, which builds on the aforementioned 2021 AI package.[116] Within the AI package instruments, the AI Act[117] adapts the rules to address the emerging risks and challenges posed by AI.[118] In addition, the package has been complemented with two new legislative instruments, namely the revised PLD[119] and the proposal for the AILD.[120]

The combination of the AI Act and these directives aims at establishing a comprehensive liability framework for AI that encompasses both preventive and compensatory measures. With the AI Act on the side of ex ante regulation, and the combination of the revised PLD and the new AILD on the side of ex post regulation, we end up with a bipartite model. In the following sections, after having provided an overview of the AI Act, we focus on the liability directives. The latter introduce sets of rules that, by encompassing different types of liability tackling different harms, aim to be truly complementary and to build an overall effective liability regime that pushes for increased trust in AI, enhances legal certainty to encourage investments in innovation, and guarantees fair compensation for harm if it were to occur in spite of the preventive AI Act requirements.

3.2.1. The AI act

Harmonizing the rules that govern AI systems is deemed necessary by the EU in order to ensure the safety of such systems, foster trust and legal certainty, and establish safeguards against non-compliance. To this end, the EU AI package encompasses the AI Act,[121] setting out obligations for economic operators of certain AI systems.

The AI Act contributes to the creation of an environment of trust by minimizing the associated risks, ensuring the safety of AI systems – along with other safety rules[122] – and compelling organizations to develop and utilize these systems in ways that prevent potential damages.

To do this, it differentiates between AI systems that are prohibited because they create unacceptable risk (such as social scoring systems or systems deploying manipulative or deceptive techniques, as per Article 5)[123] from those that are high-risk ('HRAIS') (such as AI systems used for credit scoring, or recruitment and for determining access to education),[124] and those that present low or minimal risks (such as chatbots). In particular, the risks taken into consideration relate to the impact that the use of an AI system can have on fundamental rights and Union values. The regulation minimizes such risks by introducing a system of obligations depending on the level of danger that a system presents.[125]

In detail, the AI Act imposes substantive and procedural requirements for HRAIS aimed at enhancing, inter alia, accountability and transparency (through documentation and logging obligations), accuracy, fairness and safety (through human oversight, conformity-assessment and data quality obligations), and robustness (through effective cybersecurity and risk management).[126] It allocates duties between the different types of operators involved in the supply chain (design, development and deployment) of HRAIS.[127] Such obligations fall primarily on providers.[128] Under specific circumstances, a distributor, importer, deployer, other third-party, or manufacturer, shall be considered to be a provider of a high-risk AI system, and be subject to the relevant obligations.[129] Obligations also fall on deployers of high-risk AI

[115] The EU legislative framework protecting against discrimination encompasses: Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin [2000] OJ L180/22; Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation [2000] OJ L303/16; Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast) [2006] OJ L204/23; Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services [2004] OJ L373/37. In addition, as from 2025, another directive will apply: Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services [2019] OJ L151/70. AI risks intensifying gender inequalities, as stated in European Commission, 'A Union of Equality: Gender Equality Strategy 2020-2025' (Communication) COM (2020) 152 final.

[116] AI Package (n 18).

[117] AI Act (n 18).

[118] See Section 3.2.1.

[119] Revised PLD (n 19).

[120] AILD (n 20).

[121] AI Act (n 18).

[122] Sectoral safety rules include, among others, Regulation (EU) 2019/2144 of the European Parliament and of the Council of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users (...) [2019] OJ L325/1 (Vehicle General Safety Regulation); Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices (...) [2017], OJ L117/1 (Medical Device Regulation).

[123] AI Act (n 18), art 5. For more on systems presenting unacceptable risk (manipulative systems, social scoring and biometric systems), see M Veale and F Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach' (2021) 22 Computer Law Review International 97.

[124] AI Act (n 18), art 6 and annex III.

[125] J Schuett, 'Risk Management in the Artificial Intelligence Act' (2023) European Journal of Risk Regulation <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/risk-management-in-the-artificial-intelligence-act/2E4D5707E65EFB3251A76E288BA74068>.

[126] AI Act (n 18), art 9-17.

[127] This top-down framework tends to be more fit for tangible products, rather than for upstream AI services 'which can be re-used downstream in a wide variety of unforeseeable contexts' (L Edwards, 'Expert explainer: The EU AI Act proposal' (Ada Lovelace Institute 2022) <https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/>.

[128] AI Act (n 18), art 16.

[129] AI Act (n 18), art 25.

systems.[130] Similarly to what is established by the product safety regime,[131] there is the need to ensure control over dangerous products entering the EU. Consequently, importers and distributors are bound by obligations under specific circumstances,[132] with authorized representatives potentially playing a crucial role.[133] As to AI systems that are considered less risky – namely those systems meant for interaction with natural persons, emotion recognition or biometric categorization, generation and manipulation of deep fakes – transparency obligations are set out for providers and users.[134] Moreover, operators can be required to provide access to data and documentation to authorities enforcing the law-at-hand.[135]

Infringements of the requirements and obligations lead to penalties,[136] that are first established at national level by Member States, and implemented such that they are effective, proportionate, and dissuasive (as well as accounting for SMEs, including start-ups). Introducing penalties against parties for a lack of compliance with the obligations proposed by the AI Act can be viewed as a way to hold them responsible for their handling of such systems.[137] Even though penalties for non-compliance with the risk-management obligations is a way to make parties indirectly liable for the harm that the AI deployment generates, the question still arises as to who should compensate the AI-caused harm that materializes once the risk is not managed. The AI Act introduces a limited range of remedies, namely the right to lodge a complaint with a market surveillance authority and the right to explanation of individual decision-making.[138] The former ensures market surveillance and the latter provides the possibility to request information when affected by individual decision-making; neither provides avenues to recover damages for victims of harm caused by non-compliance with AI Act obligations. Therefore, holding economic operators liable implies a step further from the new rules on AI governance, a gap in private enforcement that is somewhat filled by the liability directives analysed in the following sections.[139]

3.2.2. *The revised product liability directive and.*

In contrast to the AI Act's preventive scope, the revised PLD and proposed AILD have a compensatory objective, aimed at ensuring that victims can effectively obtain compensation if they have suffered AI-related damages that occur despite the preventive measures required under the new AI Act and existing sectoral safety rules (for instance on road vehicles[140] or medical devices).[141] By providing distinct yet overlapping means for redress in the occurrence of AI-related harm, these rules aim at promoting legal certainty for businesses and public trust in AI technology.[142] While seeking to establish liability, they also facilitate the quest for information about AI systems and indirectly push for transparency in line with the *ex ante* rules analyzed above.

In particular, the aim of the revised PLD[143] is to modernize existing rules on strict liability of producers for defective products, and to ensure legal certainty, redress and fair compensation by focusing only on certain types of harm (mainly suffered by individuals), i.e. harm caused by defective products including digital and refurbished ones, so to align the product liability regime to the digital environment.[144] As a matter of fact, the original PLD, adopted in 1985, presents various gaps with respect to the increasingly digital, circular, and global economy – specifically, it does not address AI systems, and the use of any type of software embedded in products.[145] These limitations have given rise to the need for a set of new rules, to be implemented into national law, with a much broader scope of application than its predecessor.

Hence, the revised PLD introduces several changes to the previous one in relation to both its subjective and objective scope. As to the former, its application extends beyond producers, to address all the economic operators listed in Article 8 as parties that can be held liable for defective products, adopting a "layered approach".[146] This means that compensation is sought at the next layer if parties from the previous layers cannot be held liable because they are not established in the EU or cannot be identified. Liable parties will then include, beside the manufacturer of the defective product or component, its authorised representative, the importer of the product, the so-called "fulfilment service providers", or – under certain conditions – each distributor of the product, also any natural or legal person that modifies a product and thereafter makes it available on the market or puts it into service to be held liable (as a manufacturer), but only for substantial modifications undertaken "outside the [original] manufacturer's control".[147]

[130] AI Act (n 18), art 26.

[131] See Section 3.2.2.

[132] AI Act (n 18), art 23-24. Distributors, importers, users, as well as any other third-party, have additional obligations when they modify the substance or purpose of HRAIS, or when it is placed under their name or trademark. Questions arise as to what is considered a 'substantial modification' under Article 3 (23) of the AI Act.

[133] AI Act (n 18), art 22.

[134] AI Act (n 18), art 50.

[135] AI Act (n 18), art 74.

[136] AI Act (n 18), art 99.

[137] In particular, the AI Act often refers to the concept of responsibility in various provisions – including in recitals 79 and 101, and articles 23, 24, 31, 33, 47.

[138] AI Act (n 18), art 85 and art 86. The remedies introduced by the adopted text of the AI Act aim to tackle some of the criticisms raised on the initial text, see e.g. C Castets-Renard and P Besse, 'Ex ante Accountability of the AI Act: Between Certification and Standardization, in Pursuit of Fundamental Rights in the Country of Compliance' in C Castets-Renard and J Eynard (eds), *Artificial Intelligence Law: Between Sectoral Rules and Comprehensive Regime: Comparative Law Perspectives* (Bruylant 2023).

[139] See sections 3.2.2. and 3.2.3.

[140] Vehicle General Safety Regulation (n 122).

[141] Medical Device Regulation (n 122).

[142] Revised PLD (n 19), section 1.1. of the Explanatory memorandum; AILD (n 20), section 1.1. of the Explanatory memorandum.

[143] Revised PLD (n 19).

[144] European Commission, 'New liability rules on products and AI to protect consumers and foster innovation' (European Commission Press Release, 28 September 2022) <https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5807>.

[145] European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs, 'Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products – Final report' (Publications Office 2018) <https://data.europa.eu/doi/10.2873/477640>.

[146] Revised PLD (n 19), art 8. J De Bruyne, O Dheu, and C Ducuing, 'The European Commission's Approach to Extra-Contractual Liability and AI – An Evaluation of the AI Liability Directive and the Revised Product Liability Directive' (2023) 51 Computer Law & Security Review.

[147] Revised PLD (n 19), art 8.

As to the objective scope, the notion of product encompasses "all movables, even if integrated into, or inter-connected with, another movable or an immovable", so to include a direct reference to intangible items such as digital manufacturing files and software.[148] The updated PLD would therefore also cover AI systems and AI-enabled goods, thereby victims of damage caused by their deployment will be compensated without having to demonstrate fault, but rather on the basis of a product defect.

Also the notions of defect and damage are updated. Under the revised PLD, damage refers to material losses resulting from three different causes. In addition to the traditional damages for "death or personal injury" (including to psychological health), and "damage to, or destruction of, any property" except in specific cases, also the "destruction or corruption of data that are not used for professional purposes" may now require compensation.[149] The notion of defect, instead, somewhat still resembles the one adopted under the PLD, however, the non-exclusive list of circumstances to consider in this assessment is expanded. Article 7(1)(c), for instance, allows the assessment of a defect to account for "the effect on the product of its ability to continue to learn or acquire new features after it is placed on the market or put into service".[150] As such, the product liability regime seeks to be "future-proofed" against additional improvements in self-learning AI systems that would keep developing post-deployment.

Against this background, the revised PLD encompasses two provisions that are key with respect to the harm caused by AI products. In the first place, Article 9 empowers national courts to order a defendant to disclose relevant evidence, within the limits of what is "necessary and proportionate" (the meaning of which is left to the court's discretion).[151] This can be ordered at the request of a claimant, provided that they present "facts and evidence sufficient to support the plausibility of the claim for compensation." The mechanism of the disclosure order introduced by the revised PLD – and the presumption of defect in case of non-compliance (as explained below) – aim to overcome the opacity of AI systems.

In the second place, Article 10 governs the burden of proof and the conditions for claiming compensation – i.e. proving the product defectiveness, the damage suffered, and the causal link between the two.[152] In this context the revised directive establishes two rebuttable legal presumptions. First, the presumption of the product's defectiveness is triggered under certain conditions that make up three possible scenarios :[153] (*i*) when the defendant fails to disclose relevant evidence (and to comply with the aforementioned disclosure order; *ii*) when the claimant proves that the product violates (EU or national) mandatory safety requirements specifically meant to avoid this damage; or (*iii*) when the claimant shows that the damage arose from the product's "obvious malfunction" under reasonably foreseeable use or ordinary circumstances. Second, the presumption of causality, i.e. of causal link between the product's defectiveness and the damage, activates when "it has been established that the product is defective and the damage caused is of a kind typically consistent with the defect in question".[154] This can be particularly burdensome in the case of opaque, black-box AI systems, where it is difficult, if not impossible, to understand their functioning and thus their probable contribution to the damage.

In addition, the law introduces a more general presumption that is also crucial for AI products as it allows both of the above to work simultaneously. When a national court finds that "the claimant faces excessive difficulties, in particular due to the technical or scientific complexity", to establish the product's defectiveness and/or the causal link with the damage, either or both can be assumed under certain conditions.[155] This aims to ease claimants' burden of proof in complex situations where products violate safety requirements. However, for the above presumption to apply, the claimant must still prove the likelihood of the product's defectiveness and/or the likelihood of its defectiveness causing the damage. The defendant is empowered to rebut any of the Article 10 presumptions.[156] In the latter case, it can be done by contesting the existence of excessive difficulties or the mentioned likelihood.

Among the defences that are available in the revised PLD for economic operators to escape liability, they cannot invoke the software or its upgrade or update, the lack of software updates or upgrades necessary to maintain safety, or a substantial modification of the product, if they are within the manufacturer's control.[157] In fact, the revised directive recognises that software and AI systems are updatable after having been placed on the market. Such updates, upgrades, and related services can make a product defective even if it was not defective when put into circulation, thus causing harm for which compensation can be claimed. Manufacturers can even be held liable when failing to provide "software updates or upgrades necessary to maintain safety", like to overcome cybersecurity vulnerabilities. In assessing defectiveness and liability, it is relevant to see whether manufacturers have control over the product, independently of whether it has already been placed on the market (as long as it was not before the directive's transposition). Moreover, the rights conferred pursuant to the revised PLD extinguish 10 years after products have been put into circulation: this long-stop is extended in case of any "substantial modifications" to the product, likely including software updates.[158]

In summary, the revised PLD enlarges the strict liability regime to software and, therefore, AI-products, and aligns its provisions to their functioning, an extension that should surely be welcome. However, the revised PLD also raises some doubts as to the unclarity surrounding some of its terms and provisions, including the definition of "digital services" encompassed by "related services", and the definition of "software

[148] Revised PLD (n 19), art 4(1). However, in the preamble (recital 14) an exception for free and open source software is introduced, apparently for those developers who are contributors to open source and not making a profit. The intent is to encourage innovation and research, by avoiding that developers of freeware and open-source software are treated as manufacturers, who, otherwise, would be subject to liability without any expectation of profit. However, the preamble defines commercial activity by linking it to the exchange of payment or data for the software – a definition that mirrors the development of proprietary business models which significantly differ from OSS ones.

[149] Revised PLD (n 19), art 6. See C Twigg-Flesner, C Willett and J Zhang 'Guiding Principles for Updating the Product Liability Directive for the Digital Age' (2021) ELI Innovation Paper Series 8, where it is clearly stated that '[r]evisions of the notion of 'damage' could be considered to include damage to digital elements and data.'

[150] Revised PLD (n 19), art 7.

[151] Revised PLD (n 19), art 9.

[152] Revised PLD (n 19), art 10(1).

[153] Revised PLD (n 19), art 10(2).

[154] Revised PLD (n 19), art 10(3).

[155] Revised PLD (n 19), art 10(4). Cases of technical and scientific complexity are further clarified in section 1.1 of the explanatory memorandum as 'those involving pharmaceuticals, smart products or AI-enabled products (…)' (Revised PLD (n 19)).

[156] Revised PLD (n 19), art 10(5).

[157] Revised PLD (n 19), art 11.

[158] Revised PLD (n 19), art 17(1)(b).

updates or upgrades".[159] Moreover, the question of whether "modifications" include self-learning capabilities of software depends on how the concept of "manufacturer's control" will be subject to possibly divergent interpretations.[160] Also the notion of defect remains somewhat unclear, due to undefined terminology such as "reasonably foreseeable" (when referring to product (mis)use).[161] Finally, understanding what constitutes a "substantial modification" also presents its fair share of challenges.[162] The uncertainty herein depicted will be likely clarified along the way through the interpretation given by courts at national and European level. Similarly, courts will also need to learn understanding AI systems, their functioning, and their risks. This is relevant, for example, when setting a standard of conduct to apply the general standard of care to cases involving AI.[163] Assessing whether an actor's behavior compares to that of a normal and prudent person is a task that varies according to the technological state of the art.[164] For the interpretation of product liability concepts to be adequate, national courts should also be aware of sectoral safety standards, AI characteristics (such as autonomy), as well as scientific and technical frameworks of reference.[165]

3.2.3. … the proposal for a new AI liability directive

By providing rules to compensate harm caused during the use of AI systems, the proposed AILD aims to first foster innovation in the AI sector, by reducing uncertainty for players that operate in several jurisdictions, and increasing guarantees (through instruments such as the right to rebut a liability claim based on a presumption of causality). Second, it aims to increase consumers' trust when interacting with AI, by enhancing their protection, up to the "same standards of protection when harmed by AI systems as they would be if harmed under any other circumstances."[166]

In order to enable a claimant to substantiate their claims, these rules are supposed to harmonize access to information and disclosure of evidence on HRAIS, and to significantly alleviate the burden of proof for victims to access compensation. Indeed, this instrument intends to ease access to redress by removing additional evidentiary hurdles and facilitating the legal process for claimants to prove that the defendant's fault – arising from an act or omission resulting from the use of AI systems – led to damage.

In concrete, the AILD proposes a mechanism for national claims of fault-based liability for damages caused by AI systems (intentionally or negligently, as clarified in Article 1 defining the AILD's scope). To achieve this, it provides the means to prove someone's fault in the occurrence of AI-related harm, and to offer potential compensation for victims of such damage (for all types of damage and victim). This complements the revised PLD as it addresses damages that are not caused by defect,[167] and range from breaches of privacy to damages caused by safety issues, to discrimination in recruitment processes involving AI-systems.

However, unlike the revised PLD (which is based on maximum harmonization), the AILD is based on the minimum harmonization approach, which allows Member States to introduce or maintain domestic laws that ensure additional or further support to claimants when it comes to AI-caused damage, for example regarding strict liability regimes and reversals of the burden of proof.[168] Indeed, the directive provides Member States with enough leeway on the implementation of the AI liability regime into domestic law, which is specific to each jurisdiction (for instance regarding the understanding of fault and damage) and emanates from long-established legal traditions.[169] In order to avoid overstepping on national civil liability law, the proposed directive requires Member States to adopt only two main changes to its law – namely two presumptions set out in Articles 3 and 4.

The first of these presumptions is linked to the preventive measures adopted in the AI Act. As known, the AI Act introduces specific obligations for high-risk AI systems. Complementary to those obligations, Article 3 of the AILD introduces an evidence disclosure mechanism and a rebuttable presumption of non-compliance with the duty of care. Specifically, to support liability claims of persons injured by AI systems, Article 3 allows them to seek information through national courts. The latter are empowered to order that a party bound by AI Act obligations – e.g. HRAI providers (or parties bound by their obligations), HRAI users, persons in possession of relevant information, etc. – should disclose certain relevant evidence at their disposal regarding a specific HRAIS suspected of having resulted in the damage. Disclosure should be limited to the information that is necessary and proportionate to support the liability claim. The AILD also provides the possibility for a claimant to request both the disclosure and the preservation of evidence.[170] Such disclosure could be ordered by courts both during and before the initiation of proceedings on the merits, although with conditions that vary depending on the timing.[171] On the other hand, a claimant who wishes to obtain a disclosure order must first engage in every proportionate effort to gather the relevant evidence from the defendant.[172] Potential claimants can request disclosure orders before initiation of proceedings on merits only if they had first requested such disclosure to the provider

[159] Revised PLD (n 19), art 4(3) and art 4(5).

[160] De Bruyne, Dheu and Ducuing (n 146).

[161] Revised PLD (n 19), art 7.

[162] Revised PLD (n 19), art 8(2).

[163] De Bruyne, Dheu and Ducuing (n 146).

[164] J De Bruyne, E Van Gool and A Boes, 'Wat bracht 2022 en wat brengt de toekomst op het vlak van artificiële intelligentie en buitencontractuele aansprakelijkheid?' in T. Vansweevelt and B. Weyts (eds), *Verslagboek van het Ivde Interuniversitair Congres Aansprakelijkheids-en Verzekeringsrecht* (Intersentia 2022).

[165] See for example JE Baker, LN Hobart, and MG Mittelsteadt, 'AI for Judges: A Framework - CSET Policy Brief' (Center for Security and Emerging Technology 2021) 32-46, recommending initiatives aimed at fulfilling judges' competences, in addition to enhanced clarity in the legislation.

[166] European Commission, 'Questions & Answers: AI Liability Directive' (European Commission Questions and Answers, 28 September 2022) <https://ec.europa.eu/commission/presscorner/detail/en/QANDA_22_5793>.

[167] The AILD is also aligned to the AI Act: for instance, the AILD definitions (Article 2(1-3)) mirror the AI Act's definitions as for 'AI-system'(AI Act (n 18) art 3(1)),'High-risk AI system'(AI Act(n 18)art 6),and 'provider'(AIAct(n 18) art 3(3)).This confirms the proposition that they both contribute to build the EU AI liability regime. Other concepts in the AILD, however, have ceased to be aligned with the AI Act when the text of the latter was revised. Currently, the proposed AILD still refers to 'users'(AILD (n 20), art 2(4)) in line with previous drafts of the AI Act,while the now-adopted text of the AI Act refers to 'deployer'AI Act (n 18) art 3(4)).Alignment on such concepts is crucial, as discussions on the AILD are expected to be taken up again following the progress of the negotiations on the closely linked AI Act

[168] AILD (n 20) art 3-4.

[169] Bussani and Sebok (n 50).

[170] AILD (n 20), art 3(3).

[171] AILD (n 20), art 3(1) regarding potential claimants, and art 3(2) regarding claimants.

[172] AILD (n 20), art 3(2).

(or person subject to its obligations, or user) but it was refused, and if they provide enough facts and evidence to support a plausible claim (in order to avoid "fishing expeditions"). Prior disclosure would assist in the identification of relevant evidence to support the claim and of potentially liable actors (thus also eliminating those incorrectly identified, and limiting unnecessary litigation).[173]

In any case disclosure is subject to safeguards, such as limiting access to minimum and proportionate information, in order to "prevent blanket requests".[174] When assessing the proportionality of disclosure or preservation orders, the proposed AILD requires national courts to take into account the legitimate interests of all parties (among which providers and users) and confidential information.[175] When the order concerns an (alleged) confidential information or trade secret,[176] and when a party or ex-officio submits a duly motivated request, national courts are empowered to conduct the balancing exercise between disclosure/preservation and protection of secrecy, and to adopt specific measures necessary to ensure confidentiality. These measures can include, as provided by Directive 2016/943 on trade secrets, redacting sensitive portions of rulings and restricting the number of individuals granted access to some evidence.[177]

In a claim for damages, if the defendant fails to comply with a court order to disclose or preserve evidence at its disposal (for instance because it never arranged to document or preserve it), an easing of the burden of proof applies. Indeed, under the new regime, national courts can invoke a rebuttable presumption of non-compliance with the defendant's relevant duty of care (under EU or national law) that the requested evidence was meant to prove.[178] However, the presumption only applies if the court finds excessively difficult for the claimant to prove the causal link, and so long as three specific conditions are satisfied: (*i*) it must be proved by the claimant[179] or presumed by the court[180] that the defendant (or someone whose behaviour is the defendant's responsibility) has committed a relevant fault ;[181] (*ii*) it should be considered reasonably likely, in light of the circumstances, that the defendant's fault has affected the output produced by the AI system or its failure to do so; (*iii*) it is demonstrated by the claimant that the output produced by the AI system or its failure to do so gave rise to damage.

The second presumption is encompassed in Article 4 and alleviates the claimant's burden of proof by introducing a rebuttable presumption of causality that infers the causal link between the defendant's fault and the AI system's produced output (or failure to do so) that gave rise to the damage. Although its scope is limited, the presumption helps claimants in AI-related liability cases to overcome the difficulties – exacerbated by the complexity and autonomy of AI systems – faced in providing such evidence. The presumption, however, only applies if the requirement of fault (the breach of duty) is established by the court or proved by the plaintiff.

In the case of HRAIS, it varies between claims brought against producers (or a person subject to providers' obligations under the AI Act) and those brought against users. In the former case, Article 4(2) provides that the presumption applies when the provider: (*i*) fails to comply with the obligations (exhaustively enumerated in the AILD, and directly stemming from the AI Act) that relate to the quality of data (training and testing of datasets), "transparency, human oversight, system accuracy, robustness and cybersecurity"; or (*ii*) fails to take corrective actions to remedy another breach or recall a HRAIS when the fault was identified.[182] In the latter case, Article 4(3) provides[183] that it applies when the user either fails to comply with its related AI Act obligations 'to use and monitor an AI system in accordance with accompanying instructions of use or, where appropriate, suspend or interrupt its use;[184] or exposes the AI system to input data under its control which is not relevant in view of the system's intended purpose.[185]

However, the AILD also introduces limitations to the application of the rebuttable presumption of causal link that vary according to the nature of the AI used.[186] For HRAIS, the court does not apply the presumption when the defendant demonstrates that sufficient evidence and expertise is reasonably accessible for the claimant to prove the causal link – namely through the AI Act obligations related to transparency, documentation, logging and recording. This specific exception related to HRAIS can prompt defendants' compliance with such requirements.[187] For standard non-HRAIS, instead, the court only applies the presumption where it considers it "excessively difficult for the claimant to prove the causal link"[188] – which is determined based on certain AI systems' characteristics (autonomy, opacity, etc.).

In addition, there is a more general way to rebut the burden of proof. To further incentivize disclosure, the AILD provides that if the causality

[173] G Couneson, GJ Hendrix and J Bellon, 'EU – Taking responsibility for artificial intelligence: New tort liability proposals' (*DigiLinks, Linklaters*, 3 October 2022) <https://www.linklaters.com/en/insights/blogs/digilinks/2022/october/eu—taking-responsibility-for-artificial-intelligence_new-tort-liability-proposals>.

[174] AILD (n 20), section 5.3 of the Explanatory Memorandum.

[175] AILD (n 20), art 3(4).

[176] These are specifically referenced in the AI Liability Directive, under the Trade Secret Directive (EU Directive 2016/943) and national transposing legislation. See AILD (n 20), art 3(4).

[177] This mechanism closely resembles the AI Act (n 18), art 13. However, while the latter sets out a transparency requirement focused on providing information to the deployer of the AI-system, AILD (n 20), art 3 sets out the disclosure of evidence to any victim of AI harm. The AILD facilitates the process of demonstrating fault by relying on the disclosed evidence, thus supporting damage claims for compensation.

[178] AILD (n 20), art 3(5).

[179] Such proof would have to be established by the claimant according to the applicable EU law or national rules.

[180] This would follow AILD Article 3(5): AILD (n 20), art 3(5).

[181] This fault would consist in a breach of duty of care (i.e. an obligation under national or EU legislation, such as the AI Act) that is directly intended to protect against the harm that occurred.

[182] AILD (n 20), art 4(2).

[183] AILD (n 20), art 4(3).

[184] AILD (n 20), art 29.

[185] AILD (n 20), art 29(3).

[186] For AI systems used in the context of personal non-professional activities, the AILD (AILD (n 20), art 4(6)) sets out yet another differentiated regime: the causality presumption applies only when the defendant 'has materially interfered with the conditions of the operation of the AI system', or was both required and able to determine such conditions yet did not. Non-professional users of AI systems whose behaviours do not add risk are exempted from the presumption; this aims at balancing their interests with those of victims. See G Lusardi and C Darling, 'The AI Liability Directive: EU Improves Liability Protections for Those Impacted by AI' (*Technology's Legal Edge, DLA Piper*, 6 December 2022) <https://www.technologyslegaledge.com/2022/12/the-ai-liability-directive-eu-improves-liability-protections-for-those-impacted-by-ai/#page=1>.

[187] AILD (n 20), art 4(4).

[188] AILD (n 20), art 4(5).

presumption is invoked, the defendant may rebut it by demonstrating that "sufficient evidence and expertise is reasonably accessible for the claimant to prove the causal link".[189] However, it is likely difficult to provide such sufficient evidence, i.e. that the damage suffered could not have been caused by the fault (evidence of a negative fact) or that it has been caused by another factor (requiring a complete view on the facts).

In conclusion, the AILD is to welcome even though it does not shift the burden of proof on the defendant but it only alleviates it. Some scholars consider that a complete shift would be excessively burdensome, potentially put up barriers to innovation[190] and to the adoption of AI-systems, and increase contentiousness and litigation against several potentially-liable parties.[191]

Despite the criticism that these rules raise, the AILD seems to strike a balance between innovation and protection by promoting "responsible innovation".[192] First, the development and deployment of AI systems is still encouraged by avoiding a complete shift of the burden of proof on defendants and allowing the latter to rebut the legal presumptions – especially through compliance with ex-ante measures – encourage responsible innovation rather than discourage innovation altogether. Second, the provisions requiring evidence disclosure and introducing presumptions are likely to increase transparency and strengthen the protection of victims of black-box AI systems that would otherwise not be able to prove fault or causality.

However, the approach toward an alleviated claimants' burden of proof is somewhat criticized. While the ex-ante part of the AI liability framework is very articulated, the ex-post part would not be effective enough as it still puts some heavy burden of proof (and some vulnerability) on claimants. Indeed, for the directives' presumptions to apply, claimants still have to prove numerous elements such as defect (revised PLD) or fault (AILD) on one side, and damage on the other side, as well as the nexus between these two sides.[193] For disclosure of evidence to apply, claimants (revised PLD) or potential claimants (AILD) must first "present facts and evidence sufficient to support the plausibility of a claim".[194] Furthermore, for the AILD's causality presumption to apply in the case of HRAIS covered by the AI Act, the claimant must prove non-compliance with the AI Act requirements.[195]

Some uncertainty persists as to how the various provisions that constitute the liability framework should be coordinated. In the first place, there is a potential overlap among the different pieces of legislation, particularly the AILD and the revised PLD.[196] In principle, the fault-based regime introduced by the former should not overlay the no-fault regime of the latter. The AILD clearly states that it does not affect "any rights which an injured person may have under national rules implementing" the existing PLD,[197] which however differs significantly from the revised version. The relationship between the AILD and the revised PLD thus requires further clarification, for instance on the concept of burden of proof, as there could be scenarios where the injured party is relying on both the revised PLD and the AILD to seek compensation for harm caused by an AI system. In such cases, it is unclear how the burden of proof would be allocated between the two liability regimes. The question arises as to whether the presumption of defect established under the revised PLD would still apply, or the injured party would need to prove fault – even though alleviated – under the AILD.

Doubts also arise as to the delimitation of the two directives' respective scopes.[198] For instance, consider the case of companies which use standard or complex algorithms that do not fall under the strict definition of AI,[199] such as the algorithms often used in the financial sector. Indeed, complex algorithms (or software) may not use machine learning or other AI techniques, yet they can still generate significant risks and have unintended consequences that could harm consumers. Complex software that does not rely on AI (as per the definition of AI under the AI Act) falls under the scope of the revised PLD, but does not fall under the scope of the AILD, thereby creating confusion for potential plaintiffs as to which compensation regime to rely on. In addition, the revised PLD does encompass all software but sets aside specific provisions (on the reversal of the burden of proof) solely for software that exceeds a certain level of complexity. However, these complex algorithms can create risks of unforeseeability and opacity, independently of whether they qualify as AI as defined in the legislation[200] – a consideration with potential importance for the AI Act and the AILD, which adopt a risk-based perspective.[201]

Further uncertainty accompanies the choice of directives as legal instruments, for the unclarity and lack of precision surrounding certain notions that could be applied differently according to national law and to national courts' interpretation and discretion. This could result in legal uncertainty, especially when national traditions adopt diverging approaches, as it is often the case in tort law.[202] For example, in the AILD, notions such as "fault" and "user" are left to interpretation. There is also margin for subjectivity regarding certain requirements for the application of presumptions.[203] It is indeed unclear what constitutes "all

[189] AILD (n 20), art 4(7).

[190] R De Bruin, 'Autonomous Intelligent Cars on the European Intersection of Liability and Privacy' (2016) 7 European Journal of Risk Regulation 485.

[191] B Schütte and others, 'Damages Liability for Harm Caused by Artificial Intelligence – EU Law in Flux' (2021) Helsinki Legal Studies Research Paper No. 69. The fact that the proposed AILD does not completely shift the burden of proof has also been criticized. According to some scholars, the AILD does not provide sufficient protection, for various reasons (see Conclusion). See also, P Hacker, 'The European AI Liability Directives – Critique of a Half-Hearted Approach and Lessons for the Future' (2023) 51 Computer Law & Security Review; G Borges, 'Liability for AI Systems Under Current and Future Law' (2023) 24 Computer Law Review International 1; MN Duffourc and S Gerke, 'The proposed EU Directives for AI liability leave worrying gaps likely to impact medical AI' (2023) 6 npj Digital Medicine <https://www.nature.com/articles/s41746-023-00823-w>.

[192] On the concept of "responsible innovation", see BJ Koops and others (n 14).

[193] AILD (n 20), art 4, and Revised PLD (n 19), art 10.

[194] AILD (n 20), art 3, and Revised PLD (n 19), art 9. De Bruyne, Dheu and Ducuing (n 146).

[195] AILD (n 20) art 4. SA Nawaz 'The Proposed EU AI Liability Rules: Ease or Burden?' (*European Law Blog*, 7 November 2022) <https://europeanlawblog.eu/2022/11/07/the-proposed-eu-ai-liability-rules-ease-or-burden/>.

[196] Hacker (n 191).

[197] AILD (n 20), art 1(3)(b).

[198] Hacker (n 191).

[199] For instance, credit scoring can be achieved through different techniques: some may rely on AI, while others may rely on complex statistical methods not based on AI.

[200] ZC Lipton 'The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery' (2018) 16(3) ACM Queue 31.

[201] Hacker (n 191).

[202] De Bruyne, Dheu and Ducuing (n 146).

[203] T Madiega, 'EU Legislation in Progress Briefing: Artificial intelligence liability directive' (European Parliament - European Parliamentary Research Service 2023).

proportionate attempts" to gather relevant evidence, which are required before disclosure of evidence can apply.[204] Eventually, as to the causality presumption, there are notions that are subject to case-by-case analysis such as the requirement of "reasonable likelihood" of fault influencing the output, or the "duty of care" in relation to the concept of fault, and the exception when "sufficient evidence and expertise is reasonably accessible".[205]

3.3. The initiatives for AI regulation in the EU financial markets: AI liability as a cumbersome absentee

Alongside the general initiatives on AI regulation and liability, EU institutions also initiated a series of policy measures specifically designed for the banking, financial and insurance market. Such acts were originally presented as part of the Digital Finance Package of September 2020,[206] which includes the – now approved – Regulation on Crypto Assets Markets (MiCA)[207] and Regulation on digital operational resilience (DORA)[208] as well as a proposed Regulation on distributed ledger technologies (DLT).[209] The first two, in particular, have been respectively approved in late 2022 and mid-2023, and are particularly suitable to apply to AI-based applications: as for the MiCA, for example, many transactions involving crypto-asset do occur by means of smart contract or other automated techniques.[210] On the other hand, in order to comply with the goals of the DORA, financial markets operators are required to "establish a sound network and infrastructure management using appropriate techniques, methods and protocols *including implementing automated mechanisms* to isolate affected information assets in case of cyber-attacks".[211]

The package overall aims at fostering innovation in the development of digital financial instruments while, at the same time, ensuring that AI-based products fall within the scope of financial regulation, ensuring an adequate level of protection for investors.

The DORA aims to create a uniform and homogenous framework for the security of networks and information systems of businesses and organisations operating in the financial sector, as well as for third parties providing them with Information and Communication Technologies (ICT) services like cloud computing, in order to ultimately ensure that the financial sector is able to maintain its resilience in the event of significant operational disruptions. To this scope, it consolidates and updates the requirements for cyber risks by taking into account the major risk categories of financial risk (credit risk, market risk, counterparty risk, liquidity risk, and market conduct risk) through a series of targeted qualitative interventions, concerning the capabilities of protection, detection, containment, and recovery of ICT-related accidents.[212]

As for the provisions that are most pertinent to the field of artificial intelligence, the DORA establishes a set of governance rules functional to the creation of procedures and internal control for the management of all IT risks, outlining the strategies, policies, procedures, and protocols to be adopted by the financial entity. In addition, prevention obligations are required for financial entities recurring to digital technologies for the provision of products and services: in particular, operators shall define and implement policies and procedures aimed at ensuring the resilience, continuity and availability of digital systems supporting functions that are essential for customers (e.g. credit disbursement).[213]

On the other hand, the MiCA Regulation – which was approved by the European Parliament in April 2023 – establishes a number of transparency and disclosure obligations for the issuance, public offering and admission to trading of crypto-assets on trading platforms, as well as provisions functional to the authorization and supervision of crypto-currency service providers. Additionally, the MiCA Regulation lays down specific requirements for the offer of services such as the custody and administration of crypto-assets on behalf of third parties, the execution of orders, the placement, receipt and transmission of orders on behalf of third parties, and on financial advice given on operations involving crypto-assets.

Both regulations indirectly consider liability aspects: MiCA Regulation mandates cryptocurrency issuers to behave ethically, fairly, and professionally, and explicitly states that civil liability rules should apply to crypto asset issuers and their management body ;[214] the DORA, on the contrary, does not explicitly mention liability arising from malfunctioning or errors of AI-based technologies, but puts a general responsibility rule for managing the financial entity's ICT risk on the management body of the company using such tools. Hence, as it happened with other bodies of EU financial markets law,[215] the relevant regulations refer to liability rules without, indeed, providing the normative basis for their application. This, on the one hand, strengthens the considerations regarding the structural interplay between private law and financial markets law ;[216] on the other hand, though, it implies that the liability rules on AI will have a major impact also in the financial markets, operating as a substantive means for its harmonization and development. In other words, the new rules introduced by the liability package regulating the burden of proof, as well as the identification of the potential responsible for AI-related harms, will prove themselves essential to operationalize the obligations set in financial markets law, as well as to guide financial markets actors in the organization of their supervisory and compliance strategies when recurring to AI-based techniques.

Furthermore, it should be considered that, while the prospective AI-liability framework employs a "risk-based" approach — i.e., it focuses on the characteristics of the activities performed by an AI and identifies the

[204] AILD (n 20), art 3.

[205] AILD (n 20), art 4.

[206] European Commission, 'Digital finance strategy for the EU' (Communication) COM (2020) 591 final.

[207] Regulation (EU) 2023/1114 of the European Parliament and of the Council of 31 May 2023 on markets in crypto-assets, and amending Regulations (EU) No 1093/2010 and (EU) No 1095/2010 and Directives 2013/36/EU and (EU) 2019/1937 [2023] OJ L150/40 (MiCA).

[208] Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011 [2022] OJ L333/1 (DORA).

[209] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on a pilot regime for market infrastructures based on distributed ledger technology' COM (2020) 594 final (DLT).

[210] See D Zetzsche and others, 'Remaining Regulatory Challenges in Digital Finance and Crypto-Assets after MiCA' (2023) UNSW Law Research Paper No. 23-27.

[211] DORA (n 208), art. 8 (emphasis added).

[212] DORA (n 208), recitals 22-24.

[213] DORA (n 208), art. 5-6.

[214] MiCA (n 207).

[215] Regulation (EU) No 513/2011 of the European Parliament and of the Council of 11 May 2011 amending Regulation (EC) No 1060/2009 on credit rating agencies [2011] OJ L145/30.

[216] See Section 2.

risks for the counterparty — over time, financial markets regulation has traditionally relied on a "subject-based" approach, establishing the duties of an operator based on its role in the market (e.g., financial intermediaries, etc.). Interestingly, this approach seems to stay untouched even in recent pieces of legislation related to tech-based products: for example, MiCA establishes a set of duties operating upon Crypto-Assets Service Providers (CASPs) and identifies them by operating a taxonomy of different negotiable crypto-assets. Yet, no reference is made to the specific techniques according to which crypto-asset negotiation might operate, and whether liabilities arising from MiCA should be specifically declined based on the use of AIs by CASPs. Hence, the AI-liability framework will be effective in tackling all those cases, in which a "subject-oriented" approach proves itself ineffective (such as in the case of *shadow banking* phenomena).[217]

4. Conclusion

In this article, we have analyzed the two sides of the same coin, namely the rules on AI governance and on AI liability as they both contribute to creating an AI liability framework in the EU. We have also evaluated their potential impact and interplay with regulations operating in the financial markets sector. Compliance with the obligations introduced under the AI Act is essential to enhance trust in this technology within the digital single market. Compliance efforts can include system inventories, assessment of required steps, risk assessments, and appropriate governance, meaning governance that is proportionate to the risk generated by the use of AI in the specific context. The proposed AILD and the revised PLD reinforce the importance of compliance with these obligations, particularly concerning HRAIS, and provide tools for consumers to be compensated if compliance is not enough to prevent harm. This balance between enhancing consumer trust and protection, on the one hand, and fostering innovation investment, on the other, is essential to ensure that AI benefits society. These aspects are particularly relevant considering that – as observed – the rules on AI liability will likely have an impact on regulated markets and industries as well.

The analysis of how these rules impact financial markets law has provided a privileged perspective. While sector-specific regulation often refers to notions such as liability and responsibility, in this sector the characterization and operationalization of the rules to be applied is remitted to Member States' law and, when present, to other cross-sectorial EU legislations; henceforth, harmonized liability rules on AI will play a major indirect effect in shaping how these markets will develop in the future. This does not come without any concerns: for example, while encompassing significant potential for advancing investors' protection, it has been observed that the risk-based approach characterizing the AI liability framework could be difficult to reconcile with the subject-based approach followed by financial markets law. Still, it is not to be excluded that the two systems could be properly reconciled in the future and operate together to advance investor protection. In particular, financial markets provisions, by establishing supervisory and conduct-related duties, could help avoid the aforementioned problems regarding the allocation of the burden of proof and the demonstration of causality, operating as quasi-strict liability rules based on professional diligence At the same time, liability rules could come into play whenever the deployment of AI techniques is not conducted directly under the supervision of the relevant financial market actor (e.g., when it is externalized), when it is difficult to qualify the actor as a financial market participant due to its peculiar status – a major problem, for example, in the case of shadow banking-related activities – or, more generally, in any case where financial market law fails to properly apply.

More generally, the AI liability EU framework raises further criticism in relation to the risk-based approach and its horizontal nature. First, the application of the AI Act's risk-based approach to the AILD appears highly controversial. The AILD's material scope relies on the AI Act's definition of AI systems – an alignment deemed crucial. However, the AILD is also aligned to the AI Act's risk classification of AI systems,[218] indeed some core AILD principles (i.e. disclosure of evidence and part of the burden of proof alleviation) only apply to HRAIS. Although this alignment favors coherence between intertwined pieces of legislation, applying the HRAIS categorization to AI liability regimes runs the risk of over-inclusiveness (for instance in the case of general-purpose AI systems) and under-inclusiveness (of high-risk cases).[219] The question arises as to where this under-inclusiveness arises from. The AI Act's risk-based approach focuses on the impact on society as a whole. One may argue, however, that effective liability regimes must also consider the materialization of individually pronounced risks, even when their unequal distribution among members of society places them out of the AI Act's HRAIS classification.[220] When the system's probability of damage (and thus its aggregate risk-level or expected damage) is too low to qualify it as HRAIS under the AI Act, but its risk variance between individual victims is significant, we end up with diverging classifications of social risk versus individual risk. In this case, equating risk categories in different instruments (namely, in the AI Act and the AILD) would result in the absence of effective remedies for victims of individually pronounced risks. This is the case of autonomous vehicles, which fall under the HRAIS category[221] but not under the core AI Act obligations,[222] nor – as a result – under the related AILD provisions.[223]

The application of the AI Act's risk-based approach to the AILD's core provisions also means that the latter do not cover AI practices that are prohibited under Article 5 of the AI Act (i.e. those that pose "unacceptable risks" as per the Regulation).[224] Such prohibited AI systems include, among others, subliminal AI systems, exploitative AI systems, social scoring systems and "real-time" remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement. "Non-compliance with the prohibition of the AI practices referred to in Article 5" leads to administrative fines as per Articles 99(3) and 100(2) of the AI Act.[225] However, there seem to be no liability provisions set out for these prohibited systems. The rationale appears to be that prohibited systems, by not being used in the first place, do not present any possibility for damage to occur. However, such an approach assumes that the prohibition is observed without any infringement. Instead, it could be the case that an AI-system is prohibited in theory but deployed in practice, which could also lead to damages that should be compensated. Such damages caused by prohibited AI systems would not be covered by the AILD's core provisions, including the burden of proof alleviations.

In Article 5 of the AI Act regarding prohibited AI systems, the question of harm is also of concern. The language previously used in the proposal, i.e. that of "physical or psychological harm", has been replaced

[217] Chiu and MacNeil (n 15).

[218] AILD (n 20), art 1, 2(1) and 2(2).

[219] P Hacker (n 191).

[220] P Hacker (n 191).

[221] AI Act (n 18), art 6.

[222] AI Act (n 18), art 2(2) and 112, and annex I section B. HRAIS that fall under the scope of the acts mentioned in Article 2(2), such as Automated Vehicles, are only subject to Article 112 and not to the AI Act's core obligations.

[223] AILD (n 20), art 3 and 4.

[224] AI Act (n 18), art 5.

[225] AI Act (n 18), art 99(3) and 100(2).

with the more general terminology "significant harms".[226] This comes after some commentators on the proposed AI Act had suggested to expand the scope of harm to encompass more forms of harm that were previously excluded (financial, economic, cultural harms, collective and societal harms).[227] Indeed, it had been argued that stronger protection was needed against AI-powered manipulation, to prohibit all manipulative technologies interfering with fundamental rights and leading to significant harm of all sorts (including non-subliminal manipulative AI practices, and subliminal techniques paired with target's consent). Questions remain, for instance on what is considered a significant harm.[228]

Second, the adoption of a horizontal liability regime applying to a variety of sectors does not consider the fact that AI works differently according to the sector it is used in. Indeed, it runs the risk of disregarding the intrinsic differences between distinct AI applications and the issues they raise.[229] Several voices have pointed out that "no one-size-fits-all solution can (or should) be offered" regarding liability for AI-caused damage, as it fails to recognize and overcome the challenges raised by the heterogeneity of AI uses – these voices encourage instead the implementation of a range of options, with the choice within that range to be determined by various factors".[230]

Although it might require "greater effort" from policymakers,[231] some scholars deem "more appropriate" an ad-hoc, sector-specific EU liability regime, tailored to the peculiarities of different fields (parallel to the revised PLD acting as a general rule).[232] Since liability in different sectors (e.g. transport and healthcare) needs to be regulated by different frameworks when AI is out of the picture, it might be the case also when AI is used in both sectors. Elements that would be tailored to a specific AI application could include strictly liable parties, as well as remedies and obligations of termination, non-repetition, redress, and compliance.[233] This is true, of course, for financial markets as well, as the European Parliament recently clarified by stating that in the future any "regulation of AI applications in the financial sector should consider the characteristics and market failures specific to financial markets" .[234]

In conclusion, the coordination issues and general criticism above-illustrated reverberate on the effectiveness of the EU's AI liability framework at various levels, in particular at the level of organizations and consumers. Organizations that strongly invest in AI should start accounting for this new liability framework and taking steps towards future compliance with its requirements, especially at the development or procurement stage of AI systems. Otherwise, when these systems will get deployed after the implementation deadline, the new liability regime will expose them to significant risks, ranging from reputational damage to AI liability claims. To better prepare and defend themselves from such claims against their AI-systems, organizations may consider adopting some key steps aimed at detecting potential issues and mitigating risk of litigation.[235] This could include providing training to anyone involved in the design, development, implementation and operation of AI-systems covered by the legislation. Moreover, under the AILD, providers and users of HRAIS may face the causality presumption, if they do not ensure appropriate security, monitoring, or interruption of use of the AI system when required under the AI Act. They can facilitate the rebuttal of such allegations that harm was caused by AI – and prove that it was caused by another factor – by enhancing their audit capabilities (i.e. through robust documentation and activity logs of, respectively, model testing and performance) and by conducting AI incident response planning and testing (through an AI tabletop exercise).

At the level of consumers, the liability directives aim to facilitate claims of AI-caused damage, yet it remains complicated to prove fault when it comes to complex systems. Due to the black box phenomenon, the autonomy and complexity of AI systems, it is often difficult to understand the reason (and input) behind a certain output. It is not easy to prove that the datasets used in developing a system or that the accuracy level of that system are inadequate. Logged data, for instance, can be particularly hard and costly to interpret.[236] In such cases, information about the system, as provided for by the framework, does not always do much good. If the PLD overcomes these challenges of proving fault because it provides a no-fault mechanism, such strict liability mechanism is limited to material harm[237] (in line with what many developers push for).[238] For instance, credit scoring cannot be challenged by relying on defectiveness and thus on the revised PLD, but only by proving fault under the AILD.[239] Further effort might therefore be required for the effective facilitation of redress mechanisms. In this context, some commentators push for the framework to address victims' need for specific resources – both technical and financial – in order to support their claims.[240]

In addition, it is worth pointing out also that the decision to propose an AI liability framework that amounts, as far as the ex post rules are

[226] AI Act (n 18), art 5. Amendments.

[227] N Smuha and others, 'How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act' (2021) <https://ssrn.com/abstract=3899991>.

[228] M Franklin and others, 'Missing Mechanisms of Manipulation in the EU AI Act' (2022) 35 The International FLAIRS Conference Proceedings 35 <https://doi.org/10.32473/flairs.v35i.130723>.

[229] S Whittam, 'Mind the compensation gap: towards a new European regime addressing civil liability in the age of AI' (2022) 30 International Journal of Law and Information Technology 249.

[230] Expert Group Report on Liability for Artificial Intelligence and other Emerging Digital Technologies (n 24) 36; A Bertolini, 'Artificial Intelligence and Civil Liability' (European Parliament - Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies 2020).

[231] De Bruyne, Dheu and Ducuing (n 146).

[232] Bertolini (n 230).

[233] Council of Europe Ad hoc Committee on Artificial Intelligence (CAHAI), 'Feasibility Study' CAHAI (2020) 23.

[234] See G Calzolari, 'Artificial Intelligence market and capital flows. Artificial Intelligence and the financial sector at crossroads' (European Parliament - Policy Department for Economic, Scientific and Quality of Life Policies 2021).

[235] An example of measures to take is in A Gesser and others, 'The EU AI Liability Directive Will Change Artificial Intelligence Legal Risks' (*Debevoise & Plimpton Data Blog*, 24 October 2022) <https://www.debevoise.com/insights/publications/2022/10/the-eu-ai-liability-directive-will-change>.

[236] Bertolini (n 230).

[237] Ursula Pachl, the Deputy Director of the European Consumer Organization (BEUC) has already voiced this concern while talking about the directives. See L Bertuzzi, 'The new liability rules for AI' (*The Tech Brief - Euractiv*, 30 September 2022) <https://www.euractiv.com/section/digital/podcast/the-new-liability-rules-for-ai/>.

[238] Joint Industry Letter on the PLD and AI Directive (24 August 2022) <https://ccianet.org/library/joint-industry-letter-on-the-pld-and-ai-directive/>.

[239] Nawaz (n 195).

[240] Madiega (n 203), referencing De Bruyne, Dheu and Ducuing (n 146), calling for a clearer distribution of roles, better explanation of the underlying notions and the need for technical expertise and financial resources for victims that are required to prove their claims.

concerned, to directives (AILD and revised PLD) may limit the overall harmonization effect that is pursued and confirm the fragmentation already existing among Member States as to liability regime. In particular, some commentators have also pinpointed that the mechanisms under the revised PLD and AILD require to be further harmonized, for instance by empowering potential claimants with evidence disclosure also under the revised PLD.[241] In general, the effective harmonization that comes from the adoption of the depicted AI liability regime will need to be evaluated in light of the national implementation of the directives vis-à-vis the direct application of the ex ante provisions. This is a concern that is particularly strong for the financial markets: potential gaps in the level of harmonization are expected to prove particularly problematic in financial markets, which over the years have increasingly tended (especially in the aftermath of the financial crisis) to stress the need for a maximum harmonization approach operating for all financial markets actors and aiming to ensure a level playing field, at least at the European level,[242] and subjecting them to uniform rules regardless of their location (as well exemplified by the passport mechanism).[243]

Lastly, as it has been observed,[244] the currently in-debate regulatory initiatives addressing emerging technologies in financial markets do overlook on the topic of liability, mostly focusing on prescribing disclosure and supervisory duties. At the same time, the practical implementation of the duties they introduce will have to operate in accordance with the relevant tort law provisions operating at EU and Member States' levels. For example, every time that financial markets law requires operators to ensure the robustness of the algorithmic techniques employed for the creation and delivery of financial products (e.g. as the MiFID II suitability requirements),[245] the verification of such activities will be conducted pursuant to the relevant tort law rules and standards. As a consequence, the efforts of EU institutions in the field of AI liability are likely to have a substantive impact on the harmonization of financial markets law as well, by reducing the margin of appreciation that Member State courts have in the adjudication of compensation after a misconduct to be punished under financial markets law.[246]

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

[241] Hacker (n 191).

[242] T Smith and D Geradin, 'Maintaining a level playing field when Big Tech disrupts the financial services sector' (2022), 18 European Competition Journal 129.

[243] D Busch, 'The future of EU financial law' (2022) 17 Capital Markets Law Journal 52.

[244] See Section 3.3.

[245] European Securities and Markets Authority (ESMA), 'Guidelines on certain aspects of the MiFID II suitability requirements' (2023) <https://www.esma.europa.eu/sites/default/files/2023-04/ESMA35-43-3172_Guidelines_on_certain_aspects_of_the_MiFID_II_suitability_requirements.pdf>.

[246] See F De Pascalis, 'Public Enforcement and the Civil Liability Regime in the European Regulation of Credit Rating Agencies: A Quest for Interplay' in O Cherednychenko, M Andenas (eds) *Financial Regulation and Civil Liability in European Law* (Edward Elgar Publishing 2020).