De Mari Casareto dal Verme T, 'Civil liability regimes for AI-related torts within the Italian legal system' (2024) 2 EJPLT, 172 – 190.
DOI: https://doi.org/10.57230/EJPLT242TDMCDV

EJPLT
ISSN 2704-8012



Civil liability regimes for AI-related torts within the Italian legal system

I criteri di imputazione della responsabilità civile per i danni cagionati dall'IA nell'ordinamento giuridico italiano

TOMMASO DE MARI CASARETO DAL VERME
Research Fellow in Private Law
University of Trento

Abstract

The paper addresses the issue of ascribing civil liability for torts committed by 'intelligent software agents', i.e., tangible or intangible product equipped with Artificial Intelligence (AI) in the 'modern' sense of the term. As the latest proposals and forthcoming EU legislation do not affect the core of liability regimes existing at national level, it is necessary to investigate the features and operational rules of the liability regimes destined to play a central role when it comes to AI-related torts. In the Italian legal system, the outcome of the analysis shows that the so-called 'responsibility gap' generated by AI can be filled by providing for appropriate interpretative adjustments to the 'traditional' tort law rules in the light of the peculiarities of modern AI. In this respect, Italian tort law shows to be flexible enough to also cover damages caused by smart products, still reaffirming the need to strike a balance between individuals' rights protection and incentivizing technological innovation.

© The author(s) 2024, published by Suor Orsola Benincasa Università Editrice.
This contribution is licensed under a Creative Common Attribution 4.0 International Licence CC-BY-NC-ND, all the details on the license are available at:
https://creativecommons.org/licenses/by-nc-nd/4.0/

Abstract
Il contributo affronta il tema dell'imputazione della responsabilità civile per gli illeciti commessi da agenti artificiali intelligenti, ossia da qualsiasi prodotto, materiale o immateriale, dotato di Intelligenza Artificiale (IA) in senso "moderno". Poiché le ultime proposte dell'Unione Europea (UE) sul tema non incidono sul nucleo dei criteri di imputazione della responsabilità esistenti a livello nazionale, è necessario indagare le caratteristiche e le regole operative delle regole di responsabilità destinate a svolgere un ruolo centrale quando si tratta di illeciti cagionati dall'IA. Nell'ordinamento italiano, l'esito dell'analisi mostra che il cosiddetto "responsibility gap" generato dall'IA può essere colmato attraverso un'attenta opera di interpretazione evolutiva delle norme "tradizionali" alla luce delle peculiarità dell'IA moderna. In questo senso, il diritto italiano della responsabilità civile si dimostra sufficientemente elastico da regolare anche le fattispecie di danni causati da smart products, pur riaffermando la necessità di individuare il giusto equilibrio tra tutela dei diritti e progresso tecnologico.

Keywords: Artificial Intelligence; civil liability; defective AI systems; custody over AI systems

Summary: 1. Introduction. – 2. The 'responsibility gap' doctrine. – 3. AI and tort: the ground for a horizontal approach. The Intelligent Software Agent. – 4. The liability regimes for the Intelligent Software Agent. – 5. The liability of the manufacturer. – 5.1. The proof of defectiveness: the 'model software agent'. – 5.2. The defenses. – 6. The liability of the user: the 'custody' over the software agent. – 7. Conclusions.

1. Introduction.

Artificial Intelligence (AI) represents one of the leading actors of the so-called Fourth Industrial Revolution ('Industry 4.0').[1] Despite the birth of AI as a scientific discipline dating back to the 1950s,[2] today's society is in the heat of a new 'summer' of its development, where the advancements in the research allow AI to get novel and unprecedented features. In fact, although identifying a clear definition of Artificial Intelligence is a major challenge even within the scientific community, it is possible to understand the 'modern approach' to AI research in opposition to what is to be intended as traditional AI. Russel and

[1] The origin of the term can be traced back to its first use in 2011 when, during the Hannover Messe, Henning Kagermann, Wolf-Dieter Lukas and Wolfgang Wahlster - business, technology and industry leaders as well as advisors to the German government - released their paper entitled '*Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution*'. Industri 4.0 is mainly characterized by a fusion of physical and digital spheres which gives rise to complex cyber-physical systems. For deeper insights, see K Schwab, *The Fourth Industrial Revolution* (1st edn, Portfolio Penguin, 2017); A Ustundag, E Cevikan, *Industry 4.0: Managing the Digital Transformation* (Springer, 2018) 4-5.

[2] The father of AI is often identified in Alan Turing, who in the 1950s studied the theoretical possibility of answer the question 'can machines think?' in his famous article AM Turing, 'Computing Machinery and Intelligence' (1950) 49 Mind, 433. Subsequently, the official birth of Artificial Intelligence as a real scientific discipline is traced back to 1956, the year of the *Dartmouth Summer Research Project on Artificial Intelligence*, where some scientist proposed 'that a 2 month, 10 men study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire' (J McCarthy, M Minsky, N Rochester, C Shannon, 'A Proposal for the Dartmouth Summer Research Proect on Artificial Intelligence', 31 August 1955, available at: http://www.aaai.org./ojs/index.php/aimagazine/article/view/1904).

Norvig start from a central question: 'why would we want a machine to learn?', and they provide for two main reasons: 1) the designers cannot anticipate all possible situations that the agent might find itself in; 2) sometimes human programmers have no idea how to program a solution themselves.[3]

It is possible to affirm that learning techniques represent the main feature of 'modern AI', which is capable of acting autonomously, learning from its own experience, evolving and improving its performance over time. This is possible today thanks to the implementation of machine learning (ML) and deep learning (DL) techniques combined with the huge development of computers' power and the global diffusion of Internet. In fact, the primary characteristic of modern AI lies in its ability to draw inferences and identify patterns beyond human capabilities, thanks to ML and DL algorithms that enable advanced forms of learning directly from raw data and self-generated experiences, all without relying on human expertise.[4] This marks the transition from mere automation to autonomy.[5]

However, as is often the case, the main drivers of technological innovation come with as many risks to which the law must turn its attention since they are capable of generating possible protection gaps that legislators, and the EU legislator in particular, are called upon to address.

First of all, unlike previous industrial revolutions, characterized by more or less sector-specific 'symbolic technologies', AI struggles to be classified as a proper 'technology', since by its very nature it constitutes a cross-cutting approach capable of potentially affecting any type of product to be released on the market and assuming as many facets as its potential domains of application. This is why defining AI has proved to be among the most challenging tasks for the EU in the past years. Eventually, the supranational legislator had to opt for a horizontal approach to AI regulation, resulting in the adoption of the Artificial Intelligence Act ('AI Act')[6] which sets out *ex-ante* safety rules for the release of 'high-risk' AI systems into the market, irrespective of whether they are embedded in other final products or they are placed on the market independently.[7] Within the new Regulation the definition of 'AI system' is quite broad, encompassing every «machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives,

[3] S Russel, P Norvig, *Artificial Intelligence. A modern Approach* (4th edn, Pearson Education Limited, 2022) 669.
[4] Machine learning (ML) focuses on the development of algorithms and models that enable computers to learn from data and make predictions or decisions without explicit programming. The core idea is to allow machines to improve their performance on a specific task over time by learning from experience. Deep Learning does the same but employs many layers of neural networks where the inner layers are not designed by human engineers but are learned by the system autonomously through data collection and processing, employing general learning methods. For deeper insights, see E Alpaydin, *Introduction to Machine Learning* (3rd edn, The MIT Press, 2014); I Goodfellow, Y Bengio, A Courville, *Deep Learning* (The MIT Press, 2016).
[5] D Powell, 'Autonomous Systems as Legal Agents: Directly by the Recognition of Personhood or Indirectly by the Alchemy of Algorithmic Entities' (2020) 18(1) Duke Law Technol. Rev. 310, pointing out that the distinguishing element between autonomy and mere automation is the ability to learn and adapt to unforeseen circumstances.
[6] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act, AI Act).
[7] See Article 6, par. 1 AI Act.

infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments».[8]

Second, the complexity that characterizes modern AI systems' learning capacities, on the one hand, makes it difficult (even for their creators) to understand how and why a specific output has been generated – giving rise to the 'black box' phenomenon –[9], and such challenge is even exacerbated by the information asymmetries existent between consumers and professionals who monopolize the technology and the information related to it, especially in the digital environment where consumers face a form of new and 'structural' vulnerability.[10] On the other hand, gives the algorithmic decisions an inevitable degree of unpredictability and ungovernability. In this respect, the AI Act is concerned with ensuring the highest possible level of safety for high-risk AI systems by setting preventive rules on transparency and human oversight. Nevertheless, an *ex-ante* perspective is not sufficient *per se* to prevent every technology-related risk. Therefore, there is a percentage of risk that cannot be eliminated by a preventive approach; hence a problem of *ex-post* liability arises in the case of damage caused by an AI system although regularly placed on the market.[11]

2. The 'responsibility gap' doctrine.

The problem of ascribing civil liability for damages caused by AI traces back to the so-called 'responsibility gap doctrine',[12] which highlighted the possible gaps in protection caused by the aforementioned features of modern AI systems and their possible conflict with the traditional tort rules existing mainly at national level.[13] The responsibility gap doctrine in the field of AI was

[8] Article 3(1) AI Act.

[9] On the topic, F Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press, 2015); T Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box', in T Wischmeyer, T Rademacher (eds.), *Regulating Artificial Intelligence* (Springer, 2020) 75 ff.

[10] See L Gatt, IA Caggiano, 'Consumers and Digital Environments as a Structural Vulnerability Relationship' (2022) 2 EJPLT 12, who highlight that 'the crucial point lies in the determination of a concept of vulnerability that is not linked to specific physical or psychological disabilities but is identified in the relationship between the physical person and the technological environment in which he/she operates', giving rise to an 'ontological vulnerability of human beings – in general – with respect to digital technology structures. See also L Gatt, 'Legal anthropocentrism between nature and technology: the new vulnerability of human beings' (2022) 1 EJPLT 15 ff.; AA Mollo, 'La vulnerabilità tecnologica. Neurorights ed esigenze di tutela: profili etici e giuridici' (2021) 1 EPJLT 201 ff.

[11] In the EU legal system safety rules are generally set out by preventive regulations or directives informed by the so-called 'New Legislative Framework', intended to pose minimum ex-ante safety requirements that products must comply with in order to be regularly placed on the EU market. Nevertheless, such rules can cover just a limited percentage of risks, while tort law deals with regulating also additional risks which are not entirely preventable. On these topics, see N Scannicchio, 'Il diritto privato europeo nel sistema delle fonti', in N Lipari (ed), *Trattato di diritto privato europeo*, vol I (2nd edn, Cedam 2003), 21; E Bellisario, 'Il danno da prodotto conforme tra regole preventive e regole risarcitorie' (2016) 3 Eur. dir. priv. 842; R Montinaro, 'Difetto del prodotto, norme tecniche ed informazione sui rischi di danno alla salute' (2019) 1 Resp. civ. prev. 289.

[12] A Matthias, 'The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata' (2004) 6 Eth. Inf. Tech. 175.

[13] For a general reconstruction of the debate in Italy, see C Scognamiglio, 'Responsabilità civile ed intelligenza artificiale: quali soluzioni per quali problemi?' (2023) 4 Resp. civ. prev. 1082; G Finocchiaro, 'Intelligenza artificiale, diritto e responsabilità', in P Perlingieri, S Giova, I Prisco (eds), *Rapporti civilistici e*

originally framed in the philosophical debate to indicate the concern that machine learning was likely to make it more difficult, or even impossible, to attribute 'moral' blame to individuals for the consequences of unpleasant events related to its use.[14] Following, part of tort law scholars have emphasized how the renewed degree of uncertainty and risk generated by the advent of intelligent technologies has led to the question of whether existing legal norms are adequate to regulate the phenomenon or whether there is need for their general revision.[15]

Indeed, the capacity of AI of evolving and changing autonomously – as well as the inaccessibility and uncontrollability of the information stored in it and the unpredictability of its decisions – can cause a loss of control over the system by both its programmer and user, a circumstance that could make it unjustified to ascribe responsibility to such parties. Therefore, the existence of a liability gap in cases of damages caused by autonomous and unpredictable AI systems was pointed out since control and predictability are supposed to be necessary requirements for the attribution of liability to a human being, regardless of the specific liability basis of reference.[16]

The mentioned responsibility gap would result in making it difficult or even impossible for the injured parties to get effective access to justice, due to the complexity of complying with the burden of proof lying on them according to their own national legislation when it comes to AI-related tort.[17] Such complexity exists both when an action is brought against the manufacturer of the AI system and against its user. In the former case, the applicability of the Product Liability Directive (PLD) –[18] as transposed in each Member State's legislation – comes into play, which holds the manufacturer liable for damages caused by its defective product and requires the injured party to prove the defect, the damage, and the causal link between defect and damage. In the latter case, user liability is basically grounded on fault, and the injured person is required to prove the causal link between the damage and the fault of the liable party. Other times, as in Italian law, there are strict liability rules suitable to help the injured party in fulfilling their burden of proof.

---

intelligenze artificiali: attività e responsabilità. Atti del 15° Convegno Nazionale 14-15-16 maggio 2020 (ESI, 2020), 27; F Caroccia, 'Ancora su responsabilità civile e uso delle intelligenze artificiali' (2022) 2 Contr. impr. 408; C Leanza, 'Intelligenza artificiale e diritto: ipotesi di responsabilità civile nel terzo millennio' (2021) 3 Resp. civ. prev. 1020; A Bertolini, 'Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules' (2013) 5(2) L. Innov. Tech. 233; G Sartor, 'Gli agenti software e la disciplina giuridica degli strumenti cognitivi' (2003) 1 Dir. inf. 55.

[14] F Santoni De Sio, G Mecacci, 'Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address Them' (2021) 34 Phil. Tech. 1057.

[15] The initial debate was divided between a *de lege lata* approach, which emphasized the sufficiency of existing legal norms to fill a gap that was therefore only be apparent, and a *de lege ferenda* approach, which looked, instead, favorably to the introduction of *ad hoc* legal rules for AI, emphasizing its characteristics of absolute novelty. For an overview, see U Ruffolo, 'Intelligenza artificiale, machine learning e responsabilità da algoritmo' (2019) 7 Giur. it. 1689; G Taddei Elmi, F Romano, 'Il robot tra ius condendum e ius conditum' (2016) 1 Inf. dir. 115; A Santosuosso, C Boscarato, F Caroleo, 'Robot e diritto: una prima ricognizione' (2012) 7-8 Nuova giur. civ. comm. 494.

[16] PM Asaro, 'The Liability Problem for Autonomous Artificial Agents' (2016) AAAI Spring Symposium Series 191; Id., 'Robots and Responsibility from a Legal Perspective' (2007) Proceedings of the IEEE.

[17] On the topic, see M Faccioli, 'La responsabilità civile per danni cagionati da sistemi di Intelligenza Artificiale nel prisma dell'onere della prova' (2024) 3 Resp. civ. prev. 950.

[18] Council Directive of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products (85/374/EEC).

The EU has been addressing the issues related to the AI responsibility gap since the adoption of the Resolution of 16 February 2017 on civil law rules on robotics, [19] where the European Parliament reached the conclusion that traditional tort rules were inadequate for facing the features of modern AI and that new ones were necessary. This approach also emerges in the Resolution of 20 October 2020,[20] where the European Parliament proposed a special liability regime allocating various degrees of responsibility to AI 'operators' based on the level of risk posed by the AI system. Thereafter, the European legislator took a more nuanced stance by adopting two directive proposals on 28 September 2022 aimed at harmonizing some liability rules in order to alleviate the burden of proof on injured parties. The first one, known as 'AI Liability Directive' ('AILD'), [21] addresses non-contractual fault-based liability claims brought before national courts in cases where damage originates from an AI. The second one is a new Product Liability Directive proposal[22] which aims at revising and replacing the current PLD in light of the latest developments in modern digital technologies, including AI. The new PLD has been officially adopted in October 2024[23].

Although both the mentioned proposals outline specific liability rules on the burden of proof, they do not affect the applicability of the ordinary liability regimes existing at national level. Moreover, the proof-alleviation mechanisms provided therein could even not be activated and the presumptions can be rebutted.[24] In particular, in the new PLD the proof of defectiveness remains unchanged in its substance, still grounding it on the so-called 'consumer expectation test'.[25] This choice confirms that large part of the subject will be left to the liability regimes as in force and interpreted by the courts at national level. Accordingly, the relationship between AI and civil liability must start from the analysis of the liability regimes existing at national level to test whether they are consistent with the features of modern AI.

[19] European Parliament Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) (2018/C 252/25).
[20] European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)).
[21] COM(2022) 496 final 2022/0303 (COD) 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).
[22] COM(2022) 495 final 2022/0302 (COD) Proposal for a Directive of the European Parliament and of the Council on liability for defective products.
[23] Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC.
[24] The weakness of the proposals was pointed out in the sense that, on the one hand, the presumptions provided for therein could easily be rebutted by the defendant and, on the other hand, the disclosure mechanism would not necessarily help the injured party, who might be overwhelmed by an information overload, where complex technical information could complicate the burden of proof. For deeper insights, see P Hacker, 'The European AI liability Directives – Critique of a Half-Hearted Approach and Lessons for the Future' (2023) 51 Com. L. Sec. Rev. 10; A Cioni, 'Nuovi pregi e vecchi difetti della proposta di direttiva sulla responsabilità da prodotto difettoso, con particolare riferimento all'onere della prova' (2023) 2 Resp. civ. prev. 656, 670.
[25] G Veldt, 'The New Product Liability Proposal – Fit for the Digital Age or in Need of Shaping Up?' (2023) 1 EuCML 26. See also J-S Borghetti, 'Taking EU Product Liability Law Seriously: How Can the Product Liability Directive Effectively Contribute to Consumer Protection?' (2023) 1 Fr. Journ. Pub. Pol. 33, 35 available at: https://ssrn.com/abstract=4502351, according to which maintaining the same definition of defective product is the only possible way to regulate the wide variety of products equipped with AI.

3. AI and tort: the ground for a horizontal approach. The Intelligent Software Agent.

The perspective adopted for the investigation of the problems described is horizontal and cross-cutting. Significantly, the object of analysis is AI as such, regardless, that is, of its specific technological application fields. This methodological choice is justified by the very nature of the phenomenon, which manifests itself with a transversal or 'ubiquitous'[26] dynamic with respect to specific application sectors and influences what can be considered reasonable and predictable beyond the single technological contexts of reference.[27] Therefore, the analysis must transcend the boundless variety of AI applications and search for common denominators to be taken as a reference in order to elaborate categories that are as unifying as possible and likely to be applied to every sector manned by AI.[28]

To do so, it is necessary to identify the transversal aspects of AI relevant from a legal point of view as potentially challenging the applicability of ordinary tort rules. In the interests of simplification in the legal analysis, we propose a unifying and abstract category encompassing all the relevant features of AI. In this respect, the category of 'Intelligent Software Agent' is proposed here, functional to encapsulate any type of smart product, whether physical or virtual, regardless of the specific context in which it operates.[29] Such an agent is capable of collecting data through sensors, processing raw data, planning and performing actions through the knowledge and information gained, evolving, learning and improving its performance when generating its outputs. The combination of these characteristics gives the software agent a degree of unpredictability and opacity, in that it allows algorithms to evolve from their initial programming in ways that are not always predictable and understandable by those who created them or those who use them.

Thus, although the agent undergoes an intensive training phase prior to its deployment, the rules under which it makes decisions are not entirely fixed at the production stage but may change in the course of its operation. In fact, if the agent relied solely on the knowledge imparted by its programmer rather than on its own perception, the character of autonomy would be lacking.[30]

---

[26] R Leoncini, M Macaluso, 'Gli effetti economici dell'intelligenza artificiale', in VV Cuocci, FP Lops, C Motti (eds), *La responsabilità civile nell'era digitale* (Cacucci, 2022), 90.

[27] See W Barfield, U Pagallo, *Advanced Introduction to Law and Artificial Intelligence* (Edward Elgar, 2020), 95, who highlight that the regulation of AI liability is particularly complex due to the fact that some sectors where AI is applied already have specific regulations in place. It is important to consider these existing regulatory frameworks that govern areas such as autonomous vehicles, drones, healthcare, financial services, and data protection. However, AI also influences what is considered reasonable and foreseeable beyond these individual sectors. Each example presents unique challenges in civil liability, but it is essential to first address the cross-cutting issues that affect all areas of AI application before focusing on the specific complexities of each sector.

[28] A Amidei, 'Intelligenza Artificiale e product liability: sviluppi del diritto dell'Unione europea' (2019) 7 Giur. it. 1718.

[29] For a more in-depth study, please refer to T De Mari Casareto dal Verme, *Intelligenza Artificiale e responsabilità civile. Uno studio sui criteri di imputazione* (Editoriale Scientifica, 2024), 113.

[30] Many proposals have been made for the definition of 'autonomy' with regard to computer systems. There are those who identify autonomy solely with the absence of human intervention in the functioning of the system and those who instead emphasizes the ability to evolve and modify one's own initial set of knowledge over time; others go so far as to define autonomy as the tension of the system towards a result that corresponds to its internal needs and representations. When it comes to civil liability, only software

These traits identify autonomy in decision-making and the ability to evolve as the main novelties compared to the past, potentially capable of challenging the validity of traditional tort law rules.[31]

In the view of the mentioned horizontal perspective, the liability regimes to be explored must be equally transversal and, accordingly, not sector-specific, since sector-specific liability regimes would not allow a unitary analysis to be conducted due to the dishomogeneity of the liability-generating factors contemplated in such cases.[32] This does not mean neglecting the importance of considering the specific interests underlying the various application areas of AI in each case, but before assuming a 'vertical' approach we deem necessary to investigate the common, horizontal profiles.[33]

4. The liability regimes for the Intelligent Software Agent.

First of all, it must be clarified that the notion of 'agent' employed here is not to be intended in a strict sense, but in a functional way. In fact, software agents must not be considered as proper agents – i.e., as conscious, self-aware entities – but as products, albeit highly sophisticated and capable of evolving. As regards tort law, all previous proposals that have attempted to give the software agent a nature other than product have been rejected by scholarship and institutions both for technical and legal reasons.

On the one hand, some scholars (and, at the beginning, the European Parliament as well)[34] have suggested to hold the AI system directly liable for damages, by creating an *ad hoc* legal personhood[35] on the model of how it happens already for corporations, which are entitled to rights and duties recognized by the legal system, although they are not natural persons.[36]

---

agents capable of learning and evolving over time are considered properly autonomous (and not merely automated). On the topic, see G Sartor, A Omicini, 'The Autonomy of Technological Systems and Responsibilities for their Use', in N Bhuta *et al.* (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (Cambridge University Press, 2016), 39; A Santosuosso, M Tomasi, *Diritto, scienza, nuove tecnologie* (Wolters Kluwer, 2021), 325; NF Frattari, 'Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale' (2020) 1 Contr. impr. 468.

[31] G Finocchiaro, 'Intelligenza Artificiale e responsabilità' (2020) 2 Contr. impr. 727.

[32] J-S Borghetti, 'How Can Artificial Intelligence Be Defective?', in S Lohsse, R Schulze, D Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things* (Nomos, 2019), 63,64.

[33] W Barfield, U Pagallo, 'Advanced Introduction' (n 25) 95. See also J Turner, *Robot Rules. Regulating Artificial Intelligence* (Springer, 2019), 218.

[34] It was initially proposed by the European Parliament in the Resolution of 16 February 2017 on civil law rules on robotics.

[35] On the topic, see G. Teubner, *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi* (ESI, 2019), Italian version edited by P Femia; Id., 'Digital Personhood? The Status of Autonomous Software Agents in Private Law' (May 11, 2018), available at SSRN: https://ssrn.com/abstract=3177096 or http://dx.doi.org/10.2139/ssrn.3177096; L Arnaudo, R Pardolesi, 'Ecce robot. Sulla responsabilità dei sistremi adulti di intelligenza artificiale' (2023) 4 Danno resp. 415 ff.; F Caroccia, 'Soggettività giuridica dei robot?', in G Alpa (ed), *Diritto e Intelligenza Artificiale* (Pacini Giuridica, 2020), 229; M Laukyte, 'Artificial Agents among Us: Should We Recognize Them as Agents Proper?' (2016) 18(4) Eth. Inf. Tech. 1; T Onida, F Romano, S Santoro, 'Agenti elettronici e rappresentanza volontaria nell'ordinamento giuridico italiano' (2003) 1-2 Inf. dir. 197; LB Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 North Car. L. Rev. 1231.

[36] Famous in the United States was the study by Curtis E.A. Karnow, who, based on the consideration that 'physical human beings are not the only entities protected by law, nor are they the only entities that have rights', proposed to establish the so-called 'electronic persona' for the artificial agent developed through a comparative judgement with legal persons. See CEA Karnow, 'The Encrypted Self: Fleshing out the Rights

Nevertheless, the elaboration of an *ad hoc* legal personality for software agents has been deemed inappropriate because, from a technical point of view, the state of the art still does not justify giving the software agent a substantive subjectivity of its own,[37] and, from a legal point of view, this would result in a fictitious way to create a separate asset as opposed to the physical persons who effectively control it, who would still be called upon to fund it.[38] Therefore, the attribution of legal personhood to software agents would only be an apparent solution that would complicate legal relations and, in essence, would not solve the complex problem of liability allocation.[39]

On the other hand, it has been suggested that a vicarious liability regime should be applied to damages caused by AI, by assimilating the software agent to an employee for whose actions the employer is accountable, based on the *respondeat superior* model,[40] or to a minor for whom the parents or guardians are responsible.[41] It has either been proposed that the software agent be regarded as an animal for which the owner is accountable.[42] Such proposals, albeit fascinating, do not appear to be feasible insofar as they all tend to attribute to the software agent a degree of subjectivity that does not pertain to it, and, on the other hand, the legal rules corresponding to such liability regimes in the Italian legal system possess special features with respect to the

---

of Electronic Personalities' (1994) 13(1) Journ. Comp. Inf. L. 1. Still in the US, Shawn Bayern demonstrated the possibility for a computer to effectively control a limited liability company, with the consequence that the same AI system could stand trial if the company was sued there. See S Bayern, 'The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems' (2015) 19 Stan. Tech. L. Rev. 93, 112 ff. In a critical sense with respect to Bayern's studies, see Turner, 'Robot Rules' (n 31), 178, 179, according to which 'making decisions on behalf of an entity does not mean having the same legal personality as the entity. A human controller of an LLC therefore does not become personally liable for the debts of the LLC and, presumably, neither does the AI'.

[37] The possibility of suing a machine will only be possible once it has acquired a legal status, which, on closer inspection, will only be possible with the attainment of so-called general AI, because only then will the artificial agent acquire the capacity for litigation, i.e. the capacity to understand not only the meaning of its own actions but also the relevant procedural events. In this respect, see W Barfield, U Pagallo, 'Advanced Introduction' (n 25), 7, 8.

[38] L Coppini, 'Robotica e intelligenza artificiale: questioni di responsabilità civile' (2018) 4 Pol. Dir. 730, 731; A Amidei, 'Robotica intelligente e responsabilità: profili e prospettive evolutive del quadro normativo europeo', in U Ruffolo (ed), *Intelligenza Artificiale e responsabilità. Responsabilità da algoritmo? A.I. e automobili selfdriving, automazione produttiva, robotizzazione medico-farmaceutica. A.I. e attività contrattuali. Le tendenze e discipline unionali* (Giuffrè, 2017), 65 ff.

[39] G Finocchiaro, 'Intelligenza Artificiale e responsabilità' (n 29), 713 ff.

[40] The suitability of such a liability regime is argued based on the relationship between software agent and user, which would be analogous to that which exists between employee and employer, justifying therefore, that the harm caused to third parties should be attributed to those who use makes use of its work and benefits from it. See A Lior, 'AI Entities as AI Agents: Artificial Intelligence Liability and the AI Respondeat Superior Analogy' (2020) 46 Mitch. Ham. L. Rev. 1043, 1067 ff.; P Cerka, J Grigiene, G Sirbikyte, 'Liability for Damages Caused by Artificial Intelligence' (2015) 31 Comp. L. Sec. Rev. 385 ff. See also A Beckers, G Teubner, *Three Liability Regimes for Artificial Intelligence. Algorithmic Actants, Hybrids, Crowds* (Hart Publishing, 2021), in which the authors suggest a form of digital vicarious liability similar to a *respondeat superior* hypothesis, in which liability is imputed the user of the system and the imputation criterion is the violation of rules of conduct by the machine.

[41] The analogy between the conduct of a software agent and that of a minor has been argued based on the ability of AI to learn from experience by recording and analyzing data from various sources, including interactions with the user, who would be comparable to a parent. See U Pagallo, *The law of Robots. Crimes, Contracts and Torts* (Springer, 2013), 128, 129; D Marino, G Tamburrini, 'Learning robots and human responsibility' (2006) 6 Int. rev. inf. eth. 46.

[42] The analogy between software agents and animals is supported on the consideration that, like animals, such entities would also be dangerous in that they are unpredictable and inherently irrational. See R Kelley, E Schaerer, M Gomez, M Nicolescu, 'Liability in robotics: an international perspective on robots as animals' (2010) 24(13) Adv. Rob. 1861; Taddei Elmi, Romano, 'Il robot tra ius condendum' (n 14) 127.

relationship existing between the agent and the liable third party, which do not make it appropriate to extend them by analogy outside the cases expressly contemplated therein.[43]

Therefore, it seems appropriate to consider software agents as products or 'things' endowed with evolving and adaptive capabilities. Accordingly, it is necessary to examine those liability regimes that deal with the active role of a product or thing in the causation of harm. Indeed, it has been pointed out that the use of AI in many application domains will cause a shift in focus from human to product, with a resulting predominance for product-focused rather than human-focused liability regimes.[44] Such phenomenon has been observed with regard to some typical sectoral fields of application of tort rules, such as motor vehicle liability[45] and medical liability,[46] where it is possible to experience a progressive erosion of human liability – i.e., that of the driver and of the physician – in favor of that of the producer and/or the person who exerts some sort of supervision on the product. Therefore, in most instances for damages caused by the AI system, its manufacturer and its user will be held liable, alternatively or jointly, based respectively on a malfunction of the system or on the duty to custody of it.

In the Italian legal system, this results in the need to assess the enduring enforceability of the rules designed to regulate such cases with respect to the features of autonomy and unpredictability of the intelligent software agent. In this view, it is necessary to reject the idea that the starting question is 'who' and 'why' someone should be liable for AI-related torts, thus neglecting the rigid dichotomy fault-strict liability. In fact, in Italian tort law there is a plurality of liability criteria, each responding to a different rationale depending on the context in which the damage occurs and having variable degrees of objectivity.[47] The inherently dynamic and evolving nature of AI requires to adhere to and to reaffirm such a flexible liability paradigm.

---

[43] Ruffolo, 'Intelligenza artificiale, machine learning' (n 14) 1698 ff.; Bertolini, 'Robots as products' (n 12) 227, 228; M Bassini, L Liguori, O Pollicino, 'Sistemi di Intelligenza Artificiale, responsabilità e accountability. Verso nuovi paradigmi?', in F Pizzetti (ed), *Intelligenza Artificiale, protezione dei dati personali e regolazione* (Giappichelli, 2018), 359 ff.

[44] G Calabresi, E Al Mureden, *Driverless cars. Intelligenza artificiale e futuro della mobilità* (il Mulino, 2021), 147 ff.

[45] It has been often pointed out that as vehicles become more autonomous, civil liability will progressively shift from the driver to the producer. In this respect, see MC Gaeta, 'Automazione e responsabilità civile automobilistica' (2016) 5 Resp. civ. prev. 1718 ff.; E Al Mureden, 'Autonomous cars e responsabilità civile tra disciplina vigente e prospettive de iure condendo' (2019) 3 Contr. impr. 911 and, more recently, MC Gaeta, 'Civil liability rules for autonomous vehicles accidents in Italy', in H Steege, IA Caggiano, MC Gaeta, B Von Bodungen (eds), *Autonomous Vehicles and Civil Liability in a Global Perspective. Liability law study across the world in relation to SAE J3016 Standard for Driving Automation* (Springer, 2024) 371; T De Mari Casareto dal Verme, 'Rischio da circolazione stradale, R.C. auto e veicoli a guida autonoma' (2023) 3 Biol. Journ. 275 ff. For a comprehensive and in-depth study of the relationship between self-driving cars and civil liability, see MC Gaeta, *Liability rules and self-driving cars: The evolution of tort law in the light of new technologies* (Editoriale Scientifica, 2019, reprint 2023).

[46] In the healthcare sector, it has been observed that the higher the autonomy of the AI system in making healthcare treatment decisions, the less the physician should be liable for damages caused by the algorithm and, therefore, the greater the role of the liability of the medical software manufacturer. For deeper insights, see M Faccioli, 'Intelligenza artificiale e responsabilità sanitaria' (2023) 3 Nuova giur. civ. comm., 732; G Votano, 'Intelligenza artificiale in ambito sanitario: il problema della responsabilità civile' (2022) 6 Danno resp., 673.

[47] On the topic, S Rodotà, *Il problema della responsabilità civile* (Giuffrè, 1967), 144 ff.; R Scognamiglio, 'Responsabilità per colpa e responsabilità oggettiva', in Aa Vv, *Studi in memoria di Andrea Torrente* (Giuffrè,

5. The liability of the manufacturer.

On the production side, the applicability of the PLD, as transposed into the Italian legal system,[48] comes into play as the main liability regime to be invoked in cases of AI-related torts. In this respect, it has been questioned whether the PLD was applicable to defective AI systems, since it was first disputed whether AI software falls within the notion of 'product' and, consequently, whether its creator could be considered to be a producer within the meaning of the Directive.[49]

We consider this question positively resolved because, on the one hand, we agree with the interpretation of the current PLD in the sense that it already covers software as a product,[50] and, on the other hand, the new version of the Directive now expressly treats software as a product to which this liability regime applies. It is possible, therefore, to assert the applicability of the PLD both to cases where the AI software is placed on the market independently and where it constitutes a component of a final tangible or intangible product. In any case, the consumer will not necessarily have to identify the producer of the specific defective component but may sue whoever presents themselves as the producer or places their trademark on the product.[51]

Nevertheless, the victim must fulfil the burden of proof set by Article 4 of the previous PLD – confirmed by Article 10 of the new Directive – according to which the injured person shall be required to prove the damage, the defect and the causal relationship between defect and damage. Such burden of proof operates regardless of the (only possible) alleviation mechanisms provided therein. In this regard, it has often been pointed out that the technical sophistication of modern AI systems would not allow the injured party to meet their burden of proof, especially with regard to the proof of defect.[52] This

1968), 1123; C Castronovo, *La nuova responsabilità civile. Regola e metafora* (Giuffrè, 1991), 8; C Salvi, *La responsabilità civile* (Giuffrè, 2019), 161 ff.

[48] The PLD was first transposed in the Italian legal system with the D.P.R. 24 May 1988, n. 224, then incorporated into the Consumer Code (D.lgs. 6 September 2005, n. 206).

[49] This issue arises from the debate about whether software qualifies as a product or as a service, in which latter case the PLD would not be applicable whenever, for example, the software is supplied on demand and made customized by the applicant and not in series. See Barfield, Pagallo, 'Advanced Introduction' (n 25), 11,12; A Zaccaria, 'La responsabilità del «produttore» di software' (1993) Contr. impr. 303 ff.; CL Williams, 'Not so Good: The Classification of Smart Goods Under UCC Article 2' (2018) 34 Georg. St. Un. L. Rev. 453.

[50] R Montinaro, 'Responsabilità da prodotto difettoso e tecnologie digitali tra soft law e hard law' (2020) 4 Pers. merc. 372; Ruffolo, 'Intelligenza artificiale, machine learning' (n 14) 1691.

[51] It was otfen pointed out that the person harmed by a smart product would face difficulties in detecting the specific subject against whom to direct a legal proceeding, due to the technical sophistication characterizing such devices that are made by many different components, and it is complex to attribute the defect to one of them. In practice, this obstacle can be bypassed since the PLD establishes a regime of joint and several liability among the different actors of the production chain, whereby the consumer is not burdened with identifying the 'real producer', but can take legal action against any of the parties who present themselves as such or who place their own trade mark or other distinctive sign on the product, thus burdening the production chain with the risk of identifying the 'weak link' of the chain. For such a conclusion, see the decision of the Court of Justice of the European Union, Case C-264/21, *Keskinäinen Vakuutusyhtiö Fennia v Koninklijke Philips NV*, 7 July 2022 (ECLI:EU:C:2022:536).

[52] On the topic, see C Scognamiglio, 'Responsabilità civile ed intelligenza artificiale' (n 12) 1084; GF Simonini, 'La responsabilità del fabbricante nei prodotti con sistemi di intelligenza artificiale' (2023) 4 Danno resp. 435 ff.; Montinaro, 'Responsabilità da prodotto difettoso' (n 46) 365; A Procida Mirabelli di Lauro, 'Le intelligenze artificiali tra responsabilità civile e sicurezza sociale', in Perlingieri, Giova, Prisco (eds), 'Rapporti civilistici e intelligenze artificiali' (n 12) 329, 330; C Amato, 'Product Liability and Product Security: Present and Future',

aspect is of central importance since the usability of the proof of the defect translates into the concrete possibility for those damaged by defective products to have access to justice, since case law identifies the product's defect as the indefectible prerequisite of producer liability.[53]

On the defendant's side, these problems result in the possibility that the harm caused by a self-modifying algorithmic decision could lead to the automatic application of the 'later-defect defense' and of the 'state-of-the-art defense', creating a form of producer immunity from liability,[54] which is not desirable for several reasons[55] and exacerbates the mentioned responsibility gap.

5.1. The proof of defectiveness: the 'model software agent'.

Within the PLD the defectiveness of the product represents the criterion for assessing manufacturer's liability and it must be proved by the claimant. According to Article 6 of the PLD (Article 7 of the new PLD), a product is defective when it does not provide the safety which a person is entitled to expect, taking all circumstances into account. As such, the EU legislator did not adhere to the U.S. model – which explicitly defined the well-known categories of manufacturing, design and information defect –[56] but outlined a unitary notion of defect based on the so-called 'consumer expectations test'. The concept of defect under the Directive functions as a general and flexible clause,[57] rooted in the public's legitimate safety expectations.

That of defect, thus, is a relational concept rather than a technical one,[58] i.e., the defectiveness of the product must be assessed in the light of the relationship between the actual functioning of the product and the safety that the public at large was reasonably entitled to expect taking all circumstances into account. In this respect, Italian case law[59] affirms that it is not required that

---

in Lohsse, Schulze, Staudenmayer (eds.), 'Liability for artificial intelligence' (n 30) 79 ff.; HR Sullivan, SJ Schweikart, 'Are Current Tort Liability Doctrines Adequate for Addressing Injury Caused by AI?' (2019) 21 AMA Journ. Eth. 160.

[53] In the case law, see e.g. Cass. civ. sez. III, 19 February 2016, nr. 3258; Trib. Sulmona, 23 October 2010, nr. 501.

[54] R Calo, 'Open Robotics' (2011) 70 Mar. L. Rev. 571.

[55] For deeper insights, see E Palmerini, 'AI Systems and the Issue of Liability in the European and National Regulatory Strategies', in P Morgan (ed), *Tort Liability and Autonomous Systems Accidents* (Edward Elgar, 2023), 73 ff.

[56] In the US legal system, the distinction between manufacturing defects, design defects and information defects has been formalised by the Restatement of Torts, Third, at § 2, published in 1997. In the former case, strict liability is imposed on the producer and the seller, without prejudice to actions of redress between the two. In the second, the risk/utility test is adopted in order to assess the producer's liability. Finally, the producer's liability for missing information and warnings should only arise if the information or warnings would have enabled the user of the product to avoid or reduce the risk of occurrence of that damage.

[57] C Scognamiglio, 'Clausole generali e linguaggio del legislatore: lo standard della ragionevolezza nel D.P.R. 24 maggio 1988 n. 224' (1992) Quadr. 65 ff.; Montinaro, 'Difetto del prodotto, norme tecniche' (n 10) 281; GF Simonini, 'L'approccio olistico nel danno da prodotto' (2018) 2 Danno resp. 138.

[58] F Cafaggi, 'La nozione di difetto ed il ruolo dell'informazione. Per l'adozione di un modello dinamico-relazionale di difetto in una prospettiva di riforma' (1995) Riv. crit. dir. priv. 447 ff. See also J Meltzer, R Freeman, S Thomson, in *Product liability in European Union: a report for the European Commission*, 2003.

[59] See Cass. civ. sez. III, 8 October 2007, nr. 20985, where the Court of Cassation noted that the injured party is not required to prove the specific defect of the product, it being sufficient to prove those facts, pertaining to its use, from which it is possible to derive an assessment of its unsafety. See also Cass. civ., sez. III, 15 February 2018, nr. 3692.

the consumer prove the specific material defect affecting the product, sometimes interpreting the defect as a failure to comply with the standards of reliability required by users or by the relevant laws on the subject.[60] In fact, according to the principle of proximity of evidence,[61] an inherent design or manufacturing defect is unlikely to be within the consumer's power of appreciation, so that the burden of proof is distributed among parties taking into account the concrete possibility for the plaintiff and the defendant to prove facts and circumstances falling within their respective spheres of action.[62] This is confirmed by the fact that expert opinion within the trial, although useful, is not essential to prove the defect of the product,[63] also considering that often the defective product that caused the damage goes destroyed in the event of the accident and that a direct analysis of it is not feasible.[64]

Ultimately, the assessment of the defectiveness of the product depends largely on the claimant's ability to demonstrate factual circumstances from whose overall evaluation the court may validly infer such character.[65] From this point of view, the assessment of the product's defectiveness before the Courts often builds on the manner in which the event occurred together with the characteristics of the product as perceived externally by the injured party in its appearance,[66] in relation to an assessment of the reasonable expectations.[67] To do so, according to the above-mentioned relational model, the defectiveness of the product can be derived by comparing the product's external characteristics and functioning with a hypothetical, abstract model of a safe product that embodies the reasonable expectations of the public at large. Where such an argument results in a divergence between the functioning of the actual product and that of the model, the defect proof will be deemed to have been satisfied.

---

[60] Cass. civ., sez. III, 29 May 2013, nr. 13458.

[61] The principle of proximity of evidence constitutes a criterion that is sometimes supplementary and sometimes an alternative to the general rule on the distribution of the burden of proof, in so far as it helps the court to seek the solution best suited to ensuring the effectiveness of judicial protection by placing the burden of proof on the person who has the easiest access to the knowledge of relevant information and facts. See R Rordorf, 'Onere della prova e vicinanza della prova' (2023) 1 Jus Civile. 11 ff.; M Franzoni, 'La «vicinanza della prova», quindi…' (2016) 2 Contr. impr. 360 ff.; G Vettori, 'Il diritto ad un rimedio effettivo nel diritto privato europeo' (2017) 3 Riv. dir. civ. 670. In case law, the principle has been enshrined in the famous judgment Cass. civ. sez. un., 30 October 2001, nr. 13533.

[62] On the application of the principle of proximity of evidence to product liability, see A Genovese, 'Responsabilità del produttore' (2013) Dig. disc. priv., sez. civ., 614; A Di Majo, 'La responsabilità per prodotti difettosi' (1989) 1 Riv. dir. civ. 34. Within the case law, see *inter alia* Trib. Cuneo, 17 December 2021, nr. 1071

[63] The results of technical expert opinion are not always sufficient to offer the unequivocal demonstration of the product's lack of the safety that the public is entitled to expect, since scientific evidence is not a proper proof, but merely an aid for the judge in assessing the facts and making a decision. On the topic, see A Fusaro, 'Responsabilità del produttore: la difficile prova del difetto' (2017) 6 Nuova giur. civ. comm. 898.

[64] F Raspagni, 'Responsabilità del produttore: ripensamenti e conferme in tema di onere della prova del difetto – il commento' (2015) 10 Danno resp. 958.

[65] G Stella, 'Causa ignota del danno derivante dall'uso del prodotto e responsabilità del produttore per prodotto difettoso' (2017) 5 Resp. civ. prev. 1457.

[66] G Alpa, M Bessone, *La responsabilità del produttore* (Giuffrè, 1999), 339.

[67] Case law and scholars highlight the role of the reasonable principle in the assessment of the defect, which in the Italian transposition of the PLD was translated in the word 'legitimately' (*legittimamente*). See *inter alia* Cass. civ., sez. III, 13 August 2015, n. 16808; U Carnevali, 'Prevenzione e risarcimento nelle direttive comunitarie sulla sicurezza dei prodotti' (2005) 1 Resp. civ. prev. 12.

When it comes to torts committed by intelligent software agents the described operational rules may function, still taking into account the peculiarities of such sophisticated products. In fact, the liability regime under the PLD presents itself sufficiently flexible to cover also the complexity of AI: the primary rationale of the Directive was to strike a balance between consumer protection and incentivizing technological innovation, and it does so by outlining the defect as a general and flexible clause rather than a technical notion. In this view, we hereby provide for a proposal for applying the above-mentioned relational model to smart products as well.

First of all, the product's ability to evolve and the unpredictability of its harmful decisions do not affect the burden of proof on the injured person. Indeed, it is not required to the claimant to demonstrate that the defect traces back to the production phase, since it lies on the defendant to prove the 'later-defect' in order to escape liability.[68]

Second, when it comes to products endowed with highly sophisticated learning capabilities, it must be ruled out that the injured party should have to prove the specific deficiency in the internal decision-making process of the algorithm, just as the 'traditional' injured parties are not required to prove the specific material defect that caused the damage. The injured party will be entitled to fulfil their burden of proof by referring to the manner in which the event occurred together with the characteristics of the product as externally perceivable, in accordance with the described principle of proximity of evidence. [69] In this view, it should be considered that one of the main characteristics of an intelligent software agent is that it makes decisions which manifest themselves as externally perceivable behaviors. Accordingly, the conduct of the product as perceived by the claimant will be subject to evaluation in order to assess its defectiveness.[70] It is therefore appropriate to overcome the assumption that the defect should only refer to the properties of the product and not also to its 'behaviour'.[71]

The assessment of defectiveness may be based on the deviation of the software agent's external behavior from the conduct ideally performed by a hypothetical 'model software agent',[72] similarly to what happens with the assessment of fault under Article 2043 of the Civil Code, which has now

[68] G Stella, 'Causa ignota del danno' (n 60) 1450 ff.

[69] Of the same opinion is Montinaro, 'Responsabilità da prodotto difettoso' (n 46) 385, who observes that when establishing the defectiveness of an IA system it is appropriate to employ the principle of proximity of evidence in order to distribute the burden of proof among the parties to the proceedings, in the direction of a system inspired by the effectiveness of protection so as not to make access to remedies impossible for the injured party.

[70] The appropriateness of electing the external behaviour of the system as a reference parameter for defectiveness is recognized by those who emphasize the need to carry out this assessment by means of a concrete and ex post judgement aimed at evaluating the misalignment of the product's behaviour with the standards that the user community is reasonably entitled to expect. In this respect, see Procida Mirabelli di Lauro, 'Le intelligenze artificiali' (n 48) 315. See also Beckers, Teubner, 'Three Liability Regimes' (n 38), who reconstruct the AI liability model on the basis of the violation of rules of conduct by the machine.

[71] In this respect, see G Borges, 'Liability of the Operator of AI Systems De Lege Ferenda', in S Lohsse, R Schulze, D Staudenmayer (eds), *Liability for AI. Munster Colloquia on EU Law and the Digital Economy VII* (Nomos, 2023) 171, observes that 'the concept of defect refers to the properties of the product, not to its behaviour'.

[72] For deeper insights, please refer to De Mari Casareto dal Verme, 'Intelligenza Artificiale e responsabilità civile' (n 27), 251 ff.

assumed an objective nature.[73] To do so, the model will reflect the reasonable expectations of the public at large with regard to an ideally safe conduct of the product. Such a model will vary according to the actual features of the product, the context in which it is employed, its relationship with the user, and so on, taking all circumstances into account and employing the 'reasonableness' principle[74] as a filter to constrain the possible overreaching of the parameter.[75]

Such a proposal does not differ very much from applying the *res ipsa loquitur* principle also to AI-related torts. This principle is already well-known in the legal environment and consists in the employment of presumptions to prove usually the fault[76] of an agent or the defectiveness of a product[77] by alleging facts that 'speak for themselves'. When it comes to a smart product, as a thing that really can speak for itself, the *res ipsa loquitur* principle may allow to infer the product's defectiveness – and, thus, the producer's liability – by demonstrating a 'negligent' conduct of the software agent which did not comply with the public's safety expectations.[78]

5.2. The defenses.

Once the injured party has met their burden of proof, it is up to the manufacturer to prove at least one of the circumstances listed in Article 11 of the PLD if they wish to escape liability. As already mentioned, the features of the intelligent software agent might create problems with two specific defenses – the 'later-defect defense' and the 'state-of-the-art defense' – since

[73] Fault-based liability in the Italian legal system is nowadays given an objective interpretation, in the sense that the assessment of fault does not look at the agent's subjective blame but is based on the objective deviation of the conduct of the liable party from the conduct of a hypothetical model agent complying with the rules of diligence required by the circumstances or by specific laws or regulations. See inter alia M Franzoni, L'illecito (Giuffrè, 2010), 175 ff.

[74] In this respect, see also MC Buiten, 'Product liability for defective AI' (2024) 57 Eur. Journ. L. Econ., 259 ff., who observes that in case of smart products 'under EU product liability law, courts will have to judge the reasonableness of the public's inaccurate expectations of AI'.

[75] The principle of reasonableness is now part of the daily work of judges, as its flexibility allow them to invoke it for emancipating judicial decisions from the syllogistic and literal application of the law and to adapt them to the demands of equity, fairness and justice of the individual case. At the same time, when the principle is explicitly enshrined in a legal norm (as it happened with product liability in the Italian Consumer Code), it contributes to constrain the possible judge's excessive discretion. On the topic, see G Perlingieri, 'Sul criterio di ragionevolezza', in G Perlingieri, A Fachechi (eds), *Ragionevolezza e proporzionalità nel diritto contemporaneo* (vol I ESI, 2017), 6; S Patti, *La ragionevolezza nel diritto civile* (Editoriale Scientifica, 2012), 18; A Ricci, *Il criterio della ragionevolezza nel diritto privato* (Cedam, 2007), 42.

[76] The principle of res ipsa loquitur represents a method of simplifying the burden of proof that originates in common law case law – with particular regard to *Byrne v. Boadle*, 159 Eng. Rep. 299, 300 (Ex. 1863) – dealing with the proof of fault in all those cases where the harmful event presents itself with such striking peculiarities as to 'speak for itself' about the negligence of the agent and to justify a reversal of the burden of proof. On the topic, see BS Markesinis, SF Deakin, *Tort Law* (Oxford University Press, 1999), 171 ff.; U Izzo, *La precauzione nella responsabilità civile* (Cedam, 2004), 131 ff. In the Italian legal system, it found particular application in cases of medical liability. See V Fineschi, 'Res ipsa loquitur: un principio in divenire nella definizione della responsabilità medica' (1989)2 Riv. it. med. leg. 419 ff. In the Italian case law, see Cass. civ., sez. III, 21 December 1978, nr. 6141; Cass. civ., sez. III, 22 January 1999, nr. 598; Trib. Napoli, 3 December 2014, nr. 15998; Trib. Roma, 17 November 2014, nr. 22866.

[77] On the applicability of the *res ipsa loquitur* principle to product liability, see Procida Mirabelli di Lauro, 'Le intelligenze artificiali' (n 48) 320; Montinaro, 'Difetto del prodotto, norme tecniche' (n 10) 284 ff.

[78] See R Casey, 'Robot Ipsa Loquitur' (2019) 108 Georg. L. Journ. 225, on the applicability of the *res ipsa loquitur* principle to torts committed by AI. In fact, not being necessary to prove the 'wrong code line', 'thanks to *res ipsa loquitur*, plaintiffs involved in automated accidents can rely on inference to establish fault, even when they lack direct insight into the system's underlying code'.

the learning and evolving capabilities of the software agent could lead to an automatic application of such exemptions, thus depriving the injured party of substantive access to justice. We hereby provide for an interpretation of the mentioned defenses able to prevent such a result.

As regards the 'later-defect defense', according to Article 11(c) of the new PLD the producer shall not be liable if they prove 'that it is probable that the defectiveness that caused the damage did not exist at the time the product was placed on the market, put into service or, in the case of a distributor, made available on the market, or that that defectiveness came into being after that moment'. In AI-related torts, if the damage was caused by a behaviour resulting from self-modification of the product after its deployment, such a defect could automatically be regarded as supervening and the producer would always escape liability.[79] Nevertheless, in the Italian legal system, this conclusion does not seem to be supported.

In light of the difficulty for the producer to supply negative proof of the absence of an initial defect, case law and doctrine interpret this exemption as meaning that the producer will have to supply positive proof that the defect was caused by circumstances external to and subsequent to the product's being put into circulation, such as its manipulation by third parties.[80] Thus, the defense cannot be interpreted as meaning that the product's learning capacity is *per se* capable of exempting producer's liability, since the latter is required to prove that the defect originated exclusively from a source external to the product and beyond the producer's control.[81] Nevertheless, under the new paragraph 2 of Article 11, by way of derogation the abovementioned letter (c), an economic operator shall not be exempted from liability where the defectiveness of a product is due to software, including software updates or upgrades, provided that the product is within the manufacturer's control.

As regards the 'state-of-the-art defense', according to Article11(e) of the PLD the producer shall not be liable if he proves 'that the objective state of scientific and technical knowledge at the time the product was placed on the market or put into service or during the period in which the product was within the manufacturer's control was not such that the defectiveness could be discovered'. When it comes to AI systems, the possible unpredictability of the harmful output could automatically exclude producer's liability,[82] since the defense is intended to limit producer's liability to defects that were foreseeable by the producer at the moment he placed the product on the market.

Nevertheless, the unpredictability of the defect must be read not in the light of the impossibility for the individual producer to discover the defect, but in an objective sense, i.e. according to the most advanced knowledge in all fields of

[79] S Chesterman, *We, the Robots? Regulating Artificial Intelligence and the Limits of the law* (Cambridge University Press, 2021), 96, notes that with more advanced AI systems, particularly those with the ability to modify themselves, the manufacturer could invoke as contrary evidence that the defect did not exist at the time the product was put into circulation. In the same sense, M Ratti, 'Riflessioni in materia di responsabilità civile e danno cagionato da dispositivo intelligente alla luce dell'attuale scenario normativo' (2020) 3 Contr. impr. 1174.

[80] G Stella, 'Causa ignota del danno' (n 60) 1449 ff.

[81] Sometimes such defense has been interpreted as similar to force majeure. See Alpa, Bessone, 'La responsabilità del produttore' (n 61) 259.

[82] Procida Mirabelli di Lauro, 'Le intelligenze artificiali' (n 48) 325.

science and technology.[83] Thus, the exemption only covers risks that are absolutely unforeseeable because they are absolutely unknown.[84] Therefore, the mere ability to learn, which is present in the product since its conception, can certainly not *per se* ground exemption from liability for the manufacturer, the latter having to prove the absolute unpredictability of the product's defective behaviour in the light of all the technical knowledge available at the time of its release. Then the applicability of the exemption will also be subject to an assessment of the behaviour of the AI system as totally deviating from what is reasonably foreseeable according to the state of the art, taking all circumstances into account. The development risk, then, must also be subsumed into the relational model, informed by the standard of what the producer could reasonably foresee. In this view, the reasonableness principle is capable of both defining and limiting producer's liability.[85]

6. The liability of the user: the 'custody' over the software agent.

On the user side, in Italian tort law the liability regime destined to take a central role in AI-related torts is that set forth in Article 2051 of the Civil Code, according to which for damage caused by a thing the keeper thereof shall be liable, unless he proves fortuitous. It is a strict liability regime designed to hold liable whoever exerts custody over a thing that causes damage to third parties, meaning that the injured person shall not prove that the custodian was in fault but only the causal relationship between the think and the damage.

When it comes to AI systems, it has been pointed out that the conditions for liability under Article 2051 would not be compatible with the characteristics of modern AI, due to the limited possibility of control over the software agent by the user caused by the former's ability to make decisions autonomously and to perform actions unforeseeable by that subject.[86] In other words, it would not be possible to exert custody over the intelligent software agent.[87]

First of all, it should be clarified whether the software agent falls within the notion of 'thing' under Article 2051, since the Civil Code does not provide for a definition of it. In case the AI software is a component of a physical object, there is no doubt that Article 2051 is applicable. Otherwise, we should conclude that a thing is not a thing every time its functioning is software-based. On the

[83] The reference goes to the Judgment of the Court (Fifth Chamber) of 29 May 1997, *Commission of the European Communities v United Kingdom of Great Britain and Northern Ireland*, C-300/95, in which the European Commission clarified that the Directive does not refer to a producer's ability to discover the defect on the basis of his own knowledge, but to an objective state of knowledge not limited to a specific field.
[84] R Montinaro, *Dubbio scientifico e responsabilità civile* (Giuffrè, 2012), 110 ff.
[85] For deeper insights, please refer to De Mari Casareto dal Verme, 'Intelligenza Artificiale e responsabilità civile' (n 27), 286 ff.
[86] G Sartor, 'Gli agenti software: nuovi soggetti del ciberdiritto?' (2016) 2 Contr. impr. 465; Bassini, Liguori, Pollicino, 'Sistemi di Intelligenza Artificiale' (n 41) 333 ff.; Matthias, 'The Responsibility Gap' (n 11) 175 ss.; JKC Kingston, 'Artificial Intelligence and Legal Liability', in M Brame, M Petridis (eds), *Research and Development in Intelligent Systems XXXIII: Incorporating Applications and Innovations in Intelligent Systems XXIV* (Springer, 2016), 269 ff.
[87] See M Franzoni, 'La responsabilità civile: una lunga storia ancora da scrivere' (2021) 4 Contr. impr. 1121, who expresses doubts about the applicability of Article 2051 to AI systems. See also V Di Gregorio, 'Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?' (2022) 1 Danno resp. 58; M Costanza, 'L'Intelligenza Artificiale e gli stilemi della responsabilità civile' (2019) 7 Giur. it. 1687.

contrary, when the software agent causes damages independently, more questions arise. In this respect, it seems that the notion of 'thing' under Article 2051 has been interpreted in a broad sense, encompassing also non-material entities.[88] Furthermore, the legal nature of software has been long debated and today is generally accepted that software is an entity susceptible of being possessed.[89] To support this conclusion, the concept of 'thing' should be interpreted in the light of a systematic reading of Article 2051 and Article 810, which is sufficiently broad to also include non-tangible realities for which, e.g., possession is admitted.

Accordingly, there are no valid reasons why not to consider software a thing susceptible of being under the custody of a natural or legal person.

Second, it seems that the notion of 'custody' under Article 2051 is compatible with the features of modern AI. In fact, the norm is interpreted as establishing a strict liability regime where the 'custody' does not refer to control over the thing, but it is a *de facto* relationship between the thing and the custodian that implies the possibility of avoiding the harmful consequences connected with it.[90] Custodian is, therefore, anyone who is in a position to decide and govern the conditions of use and maintenance of the thing.(264)[91] The rationale of the norm lies on the 'cheapest-cost-avoider' principle, regardless of the custodian's possibility of controlling the thing or of the unpredictability of its functioning. The main function of the rule, then, is precisely to secure compensation to the injured party even when the thing escapes human control and releases its power in an unforeseen direction or to an unforeseen extent.[92]

7. Conclusions.

The analysis carried out shows that Italian tort law is equipped with sufficiently flexible and dynamic tools to allow the existing rules to cover also cases of damages caused by intelligent software agents, albeit with the appropriate interpretative adjustments.

On the production side, the relational concept of defect makes it possible to reaffirm the functioning and rationale of the PLD as a compromise between consumer protection and the need not to discourage technological progress, with a view to ensuring a common market in which all economic actors can benefit from innovation. At the same time, for this to happen it is necessary to accept the probability that there is an ineliminable portion of risk corresponding to torts that, at least in the near future, may remain without

[88] The concept of 'thing' under Article 2051 is often interpreted to mean any inanimate element, whether movable or immovable, solid but also liquid or gaseous. See A Leonardi, '(L'accidentato terreno del) Danno da cose in custodia' (2018) 5 Resp. civ. prev. 1568; GG Greco, DM Pasanisi, B Ronchi, *I danni da cose in custodia* (Giuffrè, 2004), 51; P Zivis, 'Il danno cagionato dalle cose in custodia' (1989) 2 Nuova giur. civ. comm. 100.

[89] In this respect, see AC Nazzaro, 'Nuovi beni tra funzione e dogma' (2013) 4-5 Contr. impr., 1017, who observes that the dominant opinion sees intangible assets such as software, databases, know-how and websites as entities susceptible to possession.

[90] Within the case law, see *inter alia* Cass. civ, sez. III, 29 September 2017, nr. 22839; Cass. civ., sez. III, 29 July 2016, nr. 15761.

[91] Cass. civ., sez. III, 8 February 2012, nr. 1769; Cass. civ., sez. III, 19 May 2011, nr. 11016.

[92] P Trimarchi, *Rischio e responsabilità oggettiva* (Giuffrè, 1961), 196.

compensation. Indeed, it is natural that the introduction of technological innovations into society entails an inevitable period of adjustment necessary for the elaboration of legitimate user expectations. But in reality, the existence of a 'free zone', where the cost of damages remains where they occurred, is a risk that is always necessary to enable technological innovation.

On the user side, the legal system appears to be inspired by the idea of burdening the risk of damage on the subject placed in the best situation to prevent it, by allowing the injured party to sue the person to whom the harmful event can most immediately be traced back to, due to the risk factor introduced into the society by the thing over which they exert custody.