

# In Search of Adequate Principles for AI Civil Liability\*

## Abstract

This paper examines the civil liability of AI systems in EU legislation. The analysis discusses current problems with AI liability, proposals of the European Union legislator on how to solve existing problems and includes *de lege ferenda* suggestions. The observations' conclusions allow us to answer whether the legal solutions offered by the Products Liability Directive proposal and the AI Liability Directive proposal are adequate to the identified problems with the liability of AI systems.

**KEYWORDS:** artificial intelligence, AI systems, AI liability, fault-based liability, product liability, strict liability, EU legislation

IWONA GREDKA-LIGARSKA, PhD in law, University of Silesia in Katowice,  
ORCID – 0000-0002-9243-3158, e-mail: iwona.gredka-ligarska@us.edu.pl

## 1 | Introduction

The EU legislator notes the growing legal problems related to AI technology, including civil liability for damage caused by AI systems. The proof for the above is that we are preparing a legal framework adequate to the latest technologies, which have existed for several years. This means a legal framework that serves the development of innovation on the one hand, and guarantees the development of AI that is safe for people and trustworthy on

---

\* This article was financed by funds granted under the Research Excellence Initiative of the University of Silesia in Katowice. Project no. ZFIN11061022.

the other. The legal framework for AI technology will also include regulations dedicated to civil liability for damage caused by artificial intelligence.

Regarding security, a milestone in legal regulations is undoubtedly 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 (the AI Act)<sup>[1]</sup>. The AI Act legislative process has not yet been completed, but the AI Act will be the first regulation of this type not only in Europe but also in the world. In this way, the European Union took the leading position in developing new legal regulations on AI. The purpose of the AI Act is to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, marketing and use of artificial intelligence in line with the values of the Union. The Regulation pursues a number of overriding reasons of public interest, such as a high level of protection of health, safety and fundamental rights<sup>[2]</sup>. It ensures the free movement of AI-based goods and services cross-border, thus preventing Member States from imposing restrictions on the development, marketing and use of AI systems unless explicitly authorised by the Regulation<sup>[3]</sup>.

When deciding to introduce ground-breaking legal provisions in AI safety, the European Union legislator also decided to amend the law in the area of AI liability. The legal literature has long discussed new principles of AI liability. This is the case since the current legal regime of liability for defective products is not adequate to AI systems. As a consequence of increasing legal problems in that area, the EU legislator, in 28 September 2022, published draft proposals of two new legislative acts: the Directive of the European Parliament and of the Council on liability for defective products (the Products Liability Directive proposal)<sup>[4]</sup> and the Directive of

---

<sup>1</sup> See: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>.

<sup>2</sup> The law doctrine also noticed some disadvantages of the AI Act, which should be fixed. See e.g.: Nathalie A. Smuha *et. al.*, *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act* (Birmingham: University of Birmingham, 2021), 1-59; Michael Veale, Frederik Zuiderveen Borgesius, „Demystifying the Draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach” *Computer Law Review International* No. 4 (2021): 97-112.

<sup>3</sup> See: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>.

<sup>4</sup> Brussels, 28 September 2022, COM(2022) 495 final, 2022/0302(COD). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0495>.

the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), (the AI Liability Directive proposal)<sup>[5]</sup>.

The aim of this article is to trace the issue of civil liability of AI and to confront the conclusions of such an investigation with the latest legislative proposals of the EU legislator, as included in the Products Liability Directive proposal and the AI Liability Directive proposal. The analysis will be carried out within a broader context, from current problems with AI's liability, an overview of legislative proposals to recommendations *de lege ferenda*. The considerations made according to the presented model will allow to answer the question if the legal provisions contained in the Products Liability Directive proposal and the AI Liability Directive proposal are adequate to the identified problems with the liability for damage caused by AI systems.

The article focuses exclusively on the law of the European Union, although the question of AI liability is worldwide problem. Bearing in mind the global reach of AI and the international profile of companies developing AI, it would be optimal to prepare and introduce a universal model of AI's liability on a global scale. However, it is currently impossible to achieve this goal for at least two reasons. First, because the legal systems in different parts of the world are very different. There is no common, worldwide set of rules for civil liability<sup>[6]</sup>. Second, technological giants such as the US or China are competing to develop increasingly advanced AI and achieve global hegemony in this field. Such countries compete with one another on the technological and commercial level<sup>[7]</sup> and often have conflicting interests. Therefore, competition in the area of AI takes place not only between specific corporations but also between countries and regions of the world. As a result, the aspiration to develop a single worldwide model of AI's liability is, at least for the time being, unrealistic. On the other hand, it is legitimate and feasible to develop a regional model of AI's liability on

---

<sup>5</sup> Brussels, 28 September 2022, COM(2022) 496 final, 2022/0303 (COD). [https://commission.europa.eu/system/files/2022-09/1\\_1\\_197605\\_prop\\_dir\\_ai\\_en.pdf](https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf).

<sup>6</sup> Monika Jagielska, „Odpowiedzialność za sztuczną inteligencję”, [in:] *Prawo sztucznej inteligencji*, ed. Luigi Lai, Marek Świerczyński (Warsaw: C.H. Beck, 2020), 69-79.

<sup>7</sup> See Alan O. Sykes, „The Law and Economics of «Forced» Technology Transfer and Its Implications for Trade and Investment Policy (and the U.S.-China Trade War)” *Journal of Legal Analysis*, No. 1 (2021): 127-171.

the level of the European Union. Therefore, I have presented the research limited solely to European legislation in this article.

## 2 | AI's potential to become dangerous

The European lawmaker is determined to introduce high safety standards for high-risk AI systems. Artificial intelligence should undoubtedly remain under human control, and its development may not proceed at the expense of human rights. At this point, however, a question must be asked: can AI be completely safe? In other words, is it possible, by means of introducing appropriate legislation and exercising supervision, to develop AI systems that will not cause injury to humans or damage to human property? The safety of AI systems can be considered on two levels. The first level involves the special risk to people posed by AI systems. Algorithms developed without strict human control could assume control over us – humans (e.g. by manipulating and exerting excessive influence on the decisions and choices made by a human being). To avoid this, such pieces of legislation as the AI Act are necessary. However, there is also a second level of risk posed by AI, which can be referred to as inevitable risk. Even if AI systems are safe in the legal sense – conforming to the requirements laid down in the AI Act and in other legislative acts – the risk of damage still cannot be fully eliminated. The occurrence of random events, understood as future and uncertain events that cause personal injury or damage to human property, is something that humanity has had to deal with since the beginning of its existence on Earth, and the complete elimination of such events is impossible. The foregoing is confirmed by the system of (property and personal) insurance, as developed by humankind, whose origins date back to ancient China, Mesopotamia, Phoenicia, and then Greece or Rome<sup>[8]</sup>. Despite the huge civilizational development of mankind that has taken place since the times of the first pre-insurance transactions, the operation of the insurance system in its contemporary form is still necessary. This is the case as insurance risks deriving from forces of nature have been joined by new dangers, which, paradoxically, are a consequence of the progress

---

<sup>8</sup> Pietro Masci, „The History of Insurance: Risk, Uncertainty and Entrepreneurship” *Journal of the Washington Institute of China Studies*, No. 3 (2011): 25-68.

of civilization. As a result, even safe – in the legal sense – artificial intelligence will cause damage.

AI systems have the potential to become – as a result of interacting with their environment – dangerous artificial intelligence<sup>[9]</sup>. An immanent feature of AI is its ability to learn, to transform and to self-improve. In case of AI systems intended to interact with the environment in which they are expected to operate – and, in the first place, intended to interact with humans – the potential of AI is revealed, which eludes legislation such as the AI Act and which does not easily qualify as the product's defect. At the time of being put into circulation or put into use, an AI system is safe in the legal sense (in compliance with the requirements set out in legislative acts), and only has the potential to transform in future into a dangerous AI system and to cause damage. In this context, it is not predetermined whether or not such negative transformation is actually going to take place. This will depend on the interactions taking place between the AI system and humans. Consequently, it cannot be ruled out that the same AI system would not have undergone negative transformations and caused damage if it had interacted with another user. The user, by choosing inputs that trigger various branches in the program, acts interactively with the AI system. Thus, two different users, operating from the same fact base, may select different alternatives and generate different results. In the case of AI, because of the interactive nature of the system with the user and the non-linear approach to output, it may not be possible to determine exactly how an error occurs<sup>[10]</sup>. In the learning process, an AI system is programmed to independently adjust to the environment in which it operates in response to new data. When an AI system receives new data, the system processes the data, identifies patterns, and then develops and integrates the new patterns, allowing the system to test different hypotheses and to come up with new solutions<sup>[11]</sup>. Thus, unlike traditional software, where programmers specify predetermined outcomes that serve to explicitly confine the program's output to a limited set of possible solutions, AI programs

---

<sup>9</sup> Greg Swanson, „Non-Autonomous Artificial Intelligence Programs and Products Liability: How New AI Products Challenge Existing Liability Models and Pose New Financial Burdens” *Seattle University Law Review*, No. 3 (2019): 1201-1222.

<sup>10</sup> Maruerite E. Gerstner, „Liability Issues with Artificial Intelligence Software” *Santa Clara Law Review*, No. 1 (1993): 239-269.

<sup>11</sup> Swanson, „Non-Autonomous Artificial Intelligence Programs and Products Liability”, 1206.

are expressly designed to identify and develop original solutions<sup>[12]</sup>. The transformation of an AI system into a dangerous AI system – one that can cause damage – can take place even if the system has been exploited according to its intended use. Greg Swanson gave the following example:

An AI product is manufactured and purchased in its original form, state A. After interacting with the consumer, the AI product acquires and integrates new data provided by the consumer and subsequently begins refining and altering its internal processes (the reinforcement learning process) and “evolves” to state B. Continued consumer interactions will generate more new data, which the AI will continue to analyse and either integrate or discard. Barring pre-programmed restrictions or general computer processing limitations, this iterative process will cycle onward and the AI will become more and more distinct from its original form at the time of purchase<sup>[13]</sup>.

At the same time, Swanson concluded that even with the availability of programming restrictions, the nearly infinite body of potential data inputs from over hundreds of thousands of various AI-consumer interactions render any attempt at creating a “perfect” program impractical, if not entirely impossible<sup>[14]</sup>. In addition, Maruerite Gerstner highlighted that manufacturers of AI systems are not in a position to eliminate all risk, as far as AI safety is concerned, but have to determine what level of risk is acceptable so as to maximize the utility of an AI system and minimize liability. The more restrictions imposed on an AI system to improve its safety, the more limited the system’s ability to generate new solutions<sup>[15]</sup>. It should also be remembered that the high safety standards laid down in the AI Act apply only to high-risk AI systems. On the other hand, the capacity to transform in consequence of interactions with the environment may be a feature of not only those systems but of all AI systems in general. The potential of AI to transform into dangerous AI, as a result of interacting with the environment, refers to currently used AI systems, that are not yet autonomous. However, this potential will not disappear when AI systems become fully autonomous. What may change is the reason for

---

<sup>12</sup> Gerstner, „Liability Issues with Artificial Intelligence Software”, 242-243.

<sup>13</sup> Swanson, „Non-Autonomous Artificial Intelligence Programs and Products Liability”, 1203.

<sup>14</sup> Ibidem, 1207.

<sup>15</sup> Gerstner, „Liability Issues with Artificial Intelligence Software”, 241.

the transformation of a safe AI system into a dangerous AI – even if the AI system has been used according to its intended use and even if the system is fully safe in a legal sense (in compliance with the safety requirements of the applicable legislation). In the context of autonomous AI systems, the source of danger may be the capacity of such systems to make independent, fully autonomous decisions. It is emphasized that inasmuch as a characteristic feature of machine learning systems is their capacity to construct their own conclusions based on the available information, a distinctive feature of autonomous AI systems will be their capacity to make independent decisions, aimed at achieving the assumed objective<sup>[16]</sup>. Moreover, Stuart Russell draws attention to the fact that danger posed by artificial intelligence may relate to all AI systems whose capacities to reach the assumed objective are higher than human capacities, if the objective the AI system tries to achieve has been wrongly defined and is unfavourable for humans. If we add to that an AI system's autonomy in taking steps with the intention to achieve the assumed objective and the missing possibility of any effective human intervention, there is no doubt that AI systems pose serious danger to humans. Russell emphasizes that a key question is so called “objective function”. This objective function must be designed to serve humanity<sup>[17]</sup>.

---

<sup>16</sup> Marcin Rojszczak, „Prawne aspekty systemów sztucznej inteligencji. Zarys problemu”, [in:] *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, ed. Kinga Flaga-Gieruszyńska, Jacek Gołaczyński, Dariusz Szostek (Warsaw: C.H. Beck, 2019), 1-22.

<sup>17</sup> Stuart Russell proposed a solution that he calls an „assistance game”. According to this solution, a machine still needs information from a person to complete its task. It knows it does not know everything: „Inevitably, these machines will be uncertain about our objectives – after all, we are uncertain about them ourselves – but it turns out that this is a feature, not a bug (...). Uncertainty about objectives implies that machines will necessarily defer to humans: they will ask permission, they will accept correction, and they will allow themselves to be switched off. Removing the assumption that machines should have a definite objective means that we will need to tear out and replace part of the foundations of artificial intelligence – the basic definitions of what we are trying to do. That also means rebuilding a great deal of the superstructure – the accumulation of ideas and methods for actually doing AI”. See: Stuart Russell, *Human Compatible: AI and the Problem of Control* (UK, USA, Canada, Ireland, Australia, India, New Zealand, South Africa: Penguin Books, 2019), 12.

### 3 | Current problems with AI liability

Currently, in the legal system of the European Union, the liability for damage caused by AI systems is frequently qualified as liability for a defective product under the 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 (Directive 85/374/EEC)<sup>[18]</sup>. The approach to liability for AI along the lines of liability for a product is not specific to the EU. However, it is also a common practice in the USA, for example<sup>[19]</sup>. David C. Vladeck argues that:

Truly autonomous machines may be driving cars through our neighbourhoods or piloting drones that fly above our heads sooner than we think. So long as we can conceive of these machines as “agents” of some legal person (individual or virtual), our current system of products liability will be able to address the legal issues surrounding their introduction without significant modification<sup>[20]</sup>.

However, immanent characteristics of – so far semi-autonomous – AI, such as the ability to learn, to transform, to interact with the environment give rise to serious difficulties, which are already the case, in the application of the liability regime for defective products. Moreover, such difficulties can be associated with the most crucial rules of liability for defective products under Directive 85/374/EEC. It should be remembered that under Art. 1 of Directive 85/374/EEC, the producer is liable for the damage caused by a defect in his product. ‘Producer’ means the manufacturer of a finished product, the producer of any raw material or the manufacturer of a component part and any person who, by putting his name, trademark or other distinguishing feature on the product presents himself as its producer (Article 3(1) of Directive 85/374/EEC). Without prejudice to the liability

---

<sup>18</sup> <http://data.europa.eu/eli/dir/1985/374/oj>.

<sup>19</sup> For more on product liability regulations in Austria, Czech Republic, Denmark, England and Wales, France, Germany, Italy, the Netherlands, Norway, Poland, Spain, Switzerland, Canada, Israel, South Africa and United States of America, see: *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, ed. Piotr Machnikowski (Cambridge, Antwerp, Portland: Intersentia, 2016), 111-616.

<sup>20</sup> David C. Vladeck, „Machines without Principals: Liability Rules and Artificial Intelligence” *Washington Law Review*, No. 1 (2014): 150.



of the producer, any person who imports into the Community a product for sale, hire, leasing, or any form of distribution in the course of his business shall be deemed to be a producer within the meaning of this Directive and shall be responsible as a producer (Article 3(2) of Directive 85/374/EEC). Where the producer of the product cannot be identified, each supplier of the product shall be treated as its producer unless he informs the injured person, within a reasonable time, of the identity of the producer or of the person who supplied him with the product. The same shall apply, in the case of an imported product, if this product does not indicate the identity of the importer referred to in paragraph 2, even if the name of the producer is indicated (Article 3(3) of Directive 85/374/EEC). For the purpose of attributing liability, it is irrelevant whether the product is supplied to the final consumer directly by the producer or through other parties involved in the distribution chain<sup>[21]</sup>. The right to compensation is granted to any person who suffers personal injury or damage to property as a result of a defect in a product (Article 9 letter (a) & (b) of Directive 85/374/EEC). Damage to property may consist of damage to or destruction of any item other than the defective product itself, provided that the item was intended for private use or consumption and that it was actually used by the injured person for their own private use or consumption. According to Article 4 of Directive 85/374/EEC ‘the injured person shall be required to prove the damage, the defect and the causal relationship between defect and damage.’

Under the Directive 85/374/EEC, a producer’s liability for a defective product is based on risk. However, this liability is not absolute. The producer may avoid liability by invoking one of the exonerating conditions laid down in the Directive. Therefore, the producer will not incur liability under the provisions of the Directive if the producer proves one of the following conditions:

- a. that the producer has not placed the product into circulation;
- b. that, having regard to the circumstances, it is probable that the defect which caused the damage did not exist at the time when the product was put into circulation by him or that this defect came into being afterwards;

---

<sup>21</sup> Monika Jagielska, *Odpowiedzialność za produkt* (Warsaw: Wolters Kluwer, 2009), 98.

- c. that the product was neither manufactured by him for sale or any form of distribution for economic purpose nor manufactured or distributed by him in the course of his business;
- d. that the defect is due to compliance of the product with mandatory regulations issued by the public authorities;
- e. that the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered (use of state of the art evidence);
- f. in the case of a manufacturer of a component, that the defect is attributable to the design of the product in which the component has been fitted or to the instructions given by the manufacturer of the product (Article 7 of Directive 85/374/EEC).

Another exonerating condition is the fault of the injured person. Under Article 8(2) of Directive 85/374/EEC, the liability of the producer may be reduced or disallowed when, having regard to all the circumstances, the damage is caused both by a defect in the product and by the fault of the injured person or any person for whom the injured person is responsible. However, the liability of the producer shall not be reduced when the damage is caused both by a defect in product and by the act or omission of a third party for whom the injured person is not responsible (Article 8(1) of Directive 85/374/EEC). On the other hand, Directive 85/374/EEC does not provide for an exonerating condition in the form of force majeure. In spite of the above, in a guidance note relative to the Directive<sup>[22]</sup>, it was pointed out that force majeure may be invoked if such possibility follows from the legal regimes applicable in particular Member States. Directive 85/374/EEC, introduced in the 20th century and applicable for over 30 years, is unsuitable for AI technology. This Directive was drafted with movable items in mind which were of a completely different nature from modern AI. The biggest problem does not, however, lie in the too narrow definition of the product, as used in the Directive, which in a literal sense covers only movable items<sup>[23]</sup>. Artificial intelligence – understood as algorithms that are not movable items but intangible assets – cannot qualify as products

---

<sup>22</sup> See: Notes of the E.C. Commission on the individual Articles of the draft Directive, Commercial Laws of Europe 1986, no. 96.

<sup>23</sup> Product definition is included in Art. 2 of Directive 85/374/EEC, under which: „product” means all movables even if incorporated into another movable or into an immovable. „Product” includes electricity.

in the understanding of Directive 85/374/EEC. The legal doctrine is divided on this subject<sup>[24]</sup>. However, the European Court of Justice has dealt with the problem by using functional, rather than literal interpretation of the term product and by having regard to the fact that software is often incorporated into tangible items and that it forms their element. The most serious problems with the application of Directive 85/374/EEC to AI systems currently relate to the following: 1. the exonerating condition in the form of the risk inherent in the development, and 2. the need to prove a defect in the product (AI system) and the causal link between the defect and the resulting damage (the burden of proof is on the injured party). By referring first to the exonerating condition of risk inherent in development, it should be stressed that this prerequisite is not adequate to the liability for damage caused by AI systems. This follows from the fact that AI systems are not 'finished' products but systems whose immanent features are development and change. Under the current legal framework, producers of AI systems, in case of damage caused by such systems, can avoid liability by demonstrating that the state of scientific and technical knowledge at the time of putting the AI system into circulation did not allow for detection of the damage. Indeed, an AI system at the time of being put into circulation by the producer, may actually be free from any defects, and from any errors in design. However, throughout its entire lifecycle, the AI system will be subject to endless modifications; by learning, by self-improving and by acquiring experience under the influence of the environment with which the system interacts. Under the AI Act, providers of AI systems are imposed with an obligation to monitor those systems on an ongoing basis – from the time of placing them into circulation throughout their entire lifecycle – with regard to their compliance with the safety requirements laid down in the AI Act. However, this obligation covers only high-risk AI systems. Besides, it may be the case that not all potentially negative changes are identified in time, i.e. before damage is caused by a high-risk AI system. Moreover, certain transformations in the operation of the AI system may reveal their hazardous nature only at the time when the damage occurs. Finally, the eventuality cannot be ruled out that the AI system may cause damage and, at the same time, does not show any deviations from the safety requirements, and no defect in the system is identified. An AI system's feature of being subject to transformations throughout the

---

<sup>24</sup> Duncan Fairgrieve et. al., „Product Liability Directive”, [in:] *European Product Liability*, 46.

entire lifecycle should, therefore, be decisive for excluding the possibility to invoke, by a producer of such system, the prerequisite of development risk. Since development is an immanent feature of AI systems, why should a commercial entity that manufactured such system and put it into circulation be exempt from liability for the consequences of such development? The invocation of inherent risk in an increasing number of cases may lead to a situation where liability for harm caused by AI systems becomes a fiction, as providers of such systems will be able to easily avoid liability. The exonerating condition of risk inherent in development is justified only in relation to 'finished' products, to which Directive 85/374/EEC had been originally addressed. AI systems, as such, are not 'finished' products.

Another serious problem in the area of AI liability, understood as liability for a defective product, relates to the tracing, identifying and proving of a defect in the AI system that caused the damage. Identifying a defect in an AI system is extremely difficult, and as AI technology evolves, it will become more difficult with each passing year. Consequently, the injured person is in a doubly disadvantageous position. The injured person has to demonstrate a defect of the AI system that caused the damage and the causal relationship between the identified defect and the resulting damage. Moreover, one way or another, the injured person may be deprived of legal protection if the producer of the AI system invokes the risk inherent in development and demonstrates that, at the time of putting the system into circulation, the state of the art precluded identification of the defect that caused the damage. Israel Gilead emphasizes that:

[i]t may be very complex and expensive to establish that a product is defective or was produced in a negligent manner. To establish a design defect the plaintiff has to provide evidence indicating that an alternative design would have been safer without substantially derogating from the usefulness of the product (cost-benefit analysis). To establish a manufacturing defect the plaintiff has to show that something was wrong with the quality control over the production process. These issues may involve complex questions requiring expert opinions. (...) A change that may be needed (...) is to ease the burden of proof that in principle lies upon the plaintiff. New technologies are very complex, and it may be impossible or prohibitively expensive for the plaintiff to prove that his or her harm was indeed caused by the defendant's product-related negligence. Although present law already embraces powerful presumptions, favourable to plaintiffs, that shift the burden of proof to the defendant, it may be advisable to extend the scope

and the effectiveness of these presumptions in order to cope with the challenge of liability for new technologies<sup>[25]</sup>.

Although the opinion of the cited author refers to the terms of product liability applicable in Israel, the very postulate – to reverse the burden of proof – is also appropriate in the context of the legislation of the European Union, as analyzed in this article.

Within the legal framework of Directive 85/374/EEC, it is very difficult, and often even impossible, for the claimant to prove a defect in an AI system. This is the case despite the fact that, in court proceedings, such determinations are made with the involvement of specialized expert witnesses. Identification of a defect in an AI system poses a very serious challenge even to expert witnesses. This is due to the level of advancement and complication of AI systems<sup>[26]</sup>. Those systems are often a collection of algorithms originating not from one but from several providers. Development of a particular AI system is never a consequence of the work of a single person but an entire team of different specialists (analysts, knowledge engineers, programmers, program designers, developers, algorithm testers), and often even several expert teams. As a result, one person cannot possess the entire knowledge of the operation of such an AI system and cannot independently verify its operation and any possible defects. Additionally, the ‘black box effect’ makes it difficult to trace the AI’s decision-making process. For all the above reasons, court proceedings require the involvement of an increasing number of specialists which, however, does not in any way guarantee success. On the other hand, most certainly, this increases judicial costs and makes the proceedings longer. Consequently, the application of the terms of product liability to AI systems leads to a situation in which, along with the development of AI technologies, the costs of legal proceedings (including the costs of opinions by specialized expert witnesses) become increasingly higher. If the plaintiff loses the case, he or she will have to

---

<sup>25</sup> Israel Gilead, „Product Liability in Israel”, [in:] *European Product Liability*, 546-547.

<sup>26</sup> Gerstner underlines that „Without question, the software dealers are better able to detect a fault: they are in possession of the source code and they employ skilled workers who have experience in the field. With their technical sophistication, software dealers are better positioned to determine whether there are risks in using the software, whether those risks can be prevented, and what procedures are necessary to eliminate the problems”. Gerstner, „Liability Issues with Artificial Intelligence Software”, 254-255.

bear the costs. This can be a serious obstacle for injured parties, making the right of access to justice an illusion.

Therefore, it may seem that a solution of the outlined problems will be, first, annulment of the prerequisite of risk inherent to development in relation to AI systems and, second, reversal of the burden of proof. This would relieve the plaintiff of the need to prove a defect in the AI system and the causal link between the defect and the resulting damage. This could be done by adopting the legal presumption that damage caused with the involvement of an AI system – without intentional fault of the injured person – was caused by the defect of that system. The burden of proof that the damage suffered by the plaintiff was not caused by a defect of the AI system would be with the defendant (producer of the AI system). This could be demonstrated in two ways. Namely, the producer could prove that the AI system may indeed have a defect but there is no causal link between that fault and the damage suffered (the system's fault could not have been the cause of the damage suffered by the claimant). The second option open to the producer would be to demonstrate that the AI system with whose involvement the damage was caused operates completely regularly, i.e. no defect can be imputed to that system. On the other hand, the claimant would have to prove only that the damage was caused with the involvement of the AI system and the extent of that damage. The presented solution seems advantageous as it eliminates the difficulties in the application of Directive 85/374/EEC to damage caused by AI systems. However, in my opinion, this solution is not sufficient. This is due to the fact that the AI system that caused the damage might be free from any defects, both in the technical and legal sense. In consequence, damage may occur despite a lack of defect in the AI system (understood as defect of a product), and along with the development of AI technologies, such cases may become increasingly frequent. At the present time, it is assumed that the cause of a damage inflicted by an AI system is always some kind of defect of that system. That is, a mistake was made by a human responsible for designing, programming, or training the AI system<sup>[27]</sup>. However, increasing autonomy of AI systems – detachment from human decisions – can make it more and

---

<sup>27</sup> See the statement presented in the European Parliament Resolution of 20 October 2020 with recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence (2020/2014(INL)), (European Parliament's Resolution on AI civil liability). [https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html).

more difficult to defend the thesis that harmful operation of an AI system is a consequence of a specific mistake that can be identified and imputed to a human. This is the case as the damage caused by AI may result from interaction of the AI system with a human. In all those situations where an AI system causes damage despite the absence of any defect in the system, the injured person is deprived of compensation, which undermines the idea of legal protection.

With all the above in mind – in my opinion – it is legitimate to depart from the liability of AI systems as liability for a defective product. Defect of a product can only be considered when we are dealing with a finished product that is not subject to constant changes, and AI systems do not belong to such a category of products<sup>[28]</sup>. As technology develops, AI systems will increasingly evade the paradigm of product defect as it is known today. Therefore, approaching the issues of civil liability for artificial intelligence in terms of the defective product liability regime is a dead end. This obviously does not mean that producers of AI systems should not incur liability for their operation on the account that such systems – throughout their lifecycle – will be subject to constant transformations and self-improvement. Quite the contrary: producers of AI systems should incur civil liability for personal injuries and damage to human property caused by those systems. However, the terms of liability should be defined differently from the present legal framework<sup>[29]</sup>. I have specifically presented my proposals in this regard in the fifth chapter of this article.

---

<sup>28</sup> As I argued above, damage will not always be a consequence of a defect in an AI system. A particular AI system may be free from any design flaws but, as a result of interaction with its environment– and predominantly with humans – in a specific situation (X) it will react in an unpredictable way (Y) and cause damage (Z).

<sup>29</sup> As Gerstner underlines „[s]trict liability should be applied to software. A key consideration in the application of strict liability is the relative position of the victim with respect to the defendant. Applying strict liability allows the financial burden to be placed on the manufacturer and/or the vendor, the parties most able to bear the costs of the loss. The manufacturer is also in a better position to detect and correct flaws in the program, thus contributing to accident reduction. Fairness requires that compensation be provided to the innocent victim who has been financially damaged because of the injury. This compensation can be supplied by the manufacturer, who is in the better financial position relative to the victim. Further-more, the manufacturer can absorb the costs, either through insurance or price adjustments”. Gerstner, „Liability Issues with Artificial Intelligence Software”, 254-256.

## 4 | Proposed legislative changes

In the legislative process undertaken by the EU, intended to prepare and introduce legal provisions adequate to the latest technologies, one can notice an evolution of legislative projects from proposals most innovative and step into the future towards more conservative ones, focusing on optimal regulation of the current legal problems relating to AI technology. An example of this is the legislative work on civil liability for damage caused by artificial intelligence. The presented amendments evolved from a proposal to confer on autonomous, most advanced robots the status of electronic persons, through the European Parliament Resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)), (EP's AI Civil Liability Resolution), to the currently debated drafts of two directives: the Products Liability Directive proposal (PLD proposal) and the AI Liability Directive proposal (AILD proposal). In the context of withdrawal from the idea of introducing electronic personality, it is worth reminding the critical position expressed in relation to this concept in Experts' Open Letter addressed to the European Commission. Ultimately, also the EP's AI Civil Liability Resolution was not met with approval. Eventually, the European Commission proposed to repeal Directive 85/374/EEC and to introduce two new Directives instead: the Products Liability Directive and the AI Liability Directive. It is precisely these two directives that should provide the basis for a consistent system of civil liability for damage caused by AI systems. Already at this point, it should be noted that, in the Products Liability Directive, the legislator did not introduce a key, and highlighted in the third Chapter of this article, modification of the terms of liability for defective products by: 1. reversing the burden of proof and 2. setting aside the exonerating condition of development risk. In the draft proposal of the Products Liability Directive, however, legal instruments have been envisaged that are expected to facilitate the injured party's assertion of claims under the regime of liability for defective products, including AI systems. As pointed out in the explanatory memorandum of the above Directive proposal, the purpose of the new Products Liability Directive is, among others: easing the burden of proof in complex cases and easing restrictions on making claims while ensuring a fair balance between the legitimate interests of manufacturers, injured persons and consumers in general. The Products Liability Directive, like the previous Directive 85/374/EEC, provides for a strict liability regime for defective products.



In the PLD proposal, AI systems were clearly qualified as a product. The PLD proposal confirms that AI systems and AI-enabled goods are „products” and therefore fall within the Products Liability Directive’s scope, meaning that compensation is available when defective AI causes damage, without the injured person having to prove the manufacturer’s fault, just like for any other product. Additionally, the PLD proposal makes it clear that not only hardware manufacturers but also software providers and providers of digital services that affect how the product works (such as a navigation service in an autonomous vehicle) can be held liable. The PLD proposal ensures that manufacturers can be held liable for changes they make to products they have already placed on the market, including when these changes are triggered by software updates or machine learning. According to the definition contained in the PLD proposal, a product means ‘[a]ll movables, even if integrated into another movable or into an immovable. „Product” includes electricity, digital manufacturing files and software’<sup>[30]</sup>. The plaintiff will be required to prove the defectiveness of the product, the damage suffered and the causal link between the defectiveness and the damage<sup>[31]</sup>. In consequence, the EU legislator did not decide to reverse the burden of proof and explained the foregoing by the need to ensure a fair balance between the legitimate interests of manufacturers, injured persons and consumers in general<sup>[32]</sup>. A new legal instrument favourable to the claimant – provided for in the Products Liability Directive – is the claimant’s right to request that the court obliges the defendant to disclose appropriate evidence at the defendant’s disposal in relation to the defective product. The claimant’s right will obviously be correlated with the power of national courts to order the defendant to disclose such information on the defective product<sup>[33]</sup>. At the same time, it was emphasized that:

---

<sup>30</sup> Art. 4(1) of the PLD proposal.

<sup>31</sup> Art. 9(1) of the PLD proposal

<sup>32</sup> However, it was noticed that: „[I]njured persons, are, however, often at a significant disadvantage compared to manufacturers in terms of access to, and understanding of, information on how a product was produced and how it operates. This asymmetry of information can undermine the fair apportionment of risk, in particular in cases involving technical or scientific complexity”. See the PLD proposal, point (30-31), 19-20. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PCo495>.

<sup>33</sup> See Art. 8 of the PLD proposal.

[I]t is also necessary to alleviate the claimant's burden of proof provided that certain conditions are fulfilled. [...] In order to provide an incentive to comply with the obligation to disclose information, national courts should presume the defectiveness of a product where a defendant fails to comply with such an obligation. [...] In order to reinforce the close relationship between product safety rules and liability rules, non-compliance with such requirements should also result in a presumption of defectiveness<sup>[34]</sup>.

Accordingly, Article 9 of the Products Liability Directive proposal provides that:

2. The defectiveness of the product shall be presumed, where any of the following conditions are met: (a) the defendant has failed to comply with an obligation to disclose relevant evidence at its disposal pursuant to Article 8(1); (b) the claimant establishes that the product does not comply with mandatory safety requirements laid down in Union law or national law that are intended to protect against the risk of the damage that has occurred; or (c) the claimant establishes that the damage was caused by an obvious malfunction of the product during normal use or under ordinary circumstances. 3. The causal link between the defectiveness of the product and the damage shall be presumed, where it has been established that the product is defective and the damage caused is of a kind typically consistent with the defect in question. 4. Where a national court judges that the claimant faces excessive difficulties, due to technical or scientific complexity, to prove the defectiveness of the product or the causal link between its defectiveness and the damage, or both, the defectiveness of the product or causal link between its defectiveness and the damage, or both, shall be presumed where the claimant has demonstrated, on the basis of sufficiently relevant evidence, that: (a) the product contributed to the damage; and (b) it is likely that the product was defective or that its defectiveness is a likely cause of the damage, or both. [...].

The presumptions envisaged in the Products Liability Directive proposal are to considerably facilitate the claimant's pursuit of compensation, however, they do not entirely remove the burden of proof from the

---

<sup>34</sup> PLD proposal, point (33-35), 20-21. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0495>.

claimant. Even in especially difficult cases<sup>[35]</sup>, the defect of the product or the causal link between the defect of the product and the damage, or both, will only be presumed if the plaintiff has proved, on the basis of sufficiently relevant evidence, that the product contributed to the damage and that it is probable that the product was defective or that its defective nature is a probable cause of the damage, or both<sup>[36]</sup>. On the other hand, the defendant will have the right to question the existence of excessive difficulties and to question the probability that the product was defective or that the product's defect was the likely cause of the damage, or that both these circumstances have been the case<sup>[37]</sup>. The defendant will also have the right to rebut all the presumptions referred to in Art. 9 (2), (3) and (4) of the PLD proposal. For manufacturers (defendants), it might prove easier to rebut the presumptions, provided for by the EU legislator in Art. 9 of the PLD proposal, than for claimants to prove the defect of the product, the damage suffered and the causal link between the defect and the damage. Such situation may be the case even despite the obligation imposed on the defendant by the court to disclose the information on the defective product – as known to the defendant – and despite the presumptions laid down in relation to especially difficult cases. This is the case since these are manufacturers (defendants) that have specialist knowledge concerning the product, including technical and scientific information about the product's construction, its specific features and operating principles. When it comes to disclosure of such information at the court's request, one can never be certain if the information concerning the product, presented in order to comply with such request, will be complete and confirm the product's defect. The defendant (manufacturer), when requested by the court, may in fact disclose only the information about the product chosen by the manufacturer, and the court, without any special knowledge about that particular product, will not be able to verify whether the information as disclosed by the manufacturer is complete.

Moreover, in the PLD proposal – similarly as in Directive 85/374/EEC – the EU legislator laid down certain exonerating conditions, exempting the

---

<sup>35</sup> Matters in which the national court judges that the claimant faces excessive difficulties, due to technical or scientific complexity, to prove the defectiveness of the product or the causal link between its defectiveness and the damage, or both (Art. 9(4) of the PLD proposal). Such difficulties may arise irrespective of the disclosure by the defendant of information about the product.

<sup>36</sup> Art. 9(4) of the PLD proposal.

<sup>37</sup> Art. 9(4) of the PLD proposal.

defendant from liability for a defective product. Most of those conditions were already present in Directive 85/374/EEC and among them one can find the prerequisite of the state of the art. The EU legislator did not decide to abolish this solution. However, importantly enough, the PLD proposal includes a modification of that condition in relation to the provisions of Directive 85/374/EEC. That is to say, the PLD proposal provides that the manufacturer shall not incur liability for damage caused by a defective product if the manufacturer proves that:

[t]he objective state of scientific and technical knowledge at the time when the product was placed on the market, put into service or in the period in which the product was within the manufacturer's control was not such that the defectiveness could be discovered<sup>[38]</sup>.

Therefore, the modification of the exonerating condition of development risk is twofold. First, it was specified that the relevant benchmark should be the objective state of scientific and technical knowledge. Second, the manufacturer's liability was extended to the entire period in which the product is within the manufacturer's control. As a consequence, the protection of injured parties has been reinforced. However, the modification is not significant enough to remove problems that have been discussed in more detail in the Chapter entitled: Current Problems with AI Liability. It is generally known that learning and development are immanent features of artificial intelligence, and many properties of AI systems can be traced precisely to the condition of constant development. Consequently, leaving the condition of development risk among the exonerating conditions that exclude liability for damage caused by AI systems may lead to a situation in which, in practice, in many cases the manufacturer's liability will be excluded. This is because, according to the objective state of scientific and technical knowledge, the defectiveness of an AI system may not be detectable. In effect, what the EU legislator could actually achieve for the enhancement of protection of injured parties due to the expansion of the definition of product defectiveness<sup>[39]</sup> will be thwarted by leaving to manufacturers the possibility of relying on the state of the art. The above relates to the extension of the defectiveness concept to factual situations in which the product does not ensure the safety that might be expected by

---

<sup>38</sup> Art. 10(1e) of the PLD proposal.

<sup>39</sup> See Art. 6 of the PLD proposal.

the public at large taking into consideration: „[t]he effect on the product of any ability to continue to learn after deployment”<sup>[40]</sup> This means that the EU legislator has fully noticed the possibility of defectiveness of an AI system as a result of that system’s further learning; already after being placed on the market or put into service. On the other hand, however, the same legislator does not ensure full legal protection to persons injured by the product’s defectiveness as defined in Art. 6(1c) of the PLD proposal. If such defectiveness arises and leads to the emergence of damage but, at the same time, the defectiveness is undetectable according to the current state of scientific and technical knowledge, the manufacturer will be in a position to avoid liability, and the injured person will be left without compensation.

In the Products Liability Directive proposal, the principle of maximum harmonization was adopted. According to Art. 3 of the PLD proposal:

[M]ember States shall not maintain or introduce, in their national law, provisions diverging from those laid down in this Directive, including more, or less, stringent provisions to achieve a different level of consumer protection, unless otherwise provided for in this Directive.

Complementation of the civil liability regime for damages caused by artificial intelligence is to be the AILD proposed by the European Commission. The Products Liability Directive and the AI Liability Directive are expected to make up a consistent and complete system of legal protection for persons injured by AI systems. According to the assumptions of the EU legislator, the principles of civil liability for damages caused by AI systems are to be compatible with the safety principles of those systems. The AI Liability Directive will apply to non-contractual, civil law claims for compensation based on fault in situations when the damage is caused by an artificial intelligence system<sup>[41]</sup>. The harmonization envisaged in the AI Liability Directive proposal is to cover only such national provisions of law concerning fault-based liability that govern the burden of proof in relation to

---

<sup>40</sup> Art. 6(1c) of the PLD proposal.

<sup>41</sup> In Art. 1(1) of the AILD, it is indicated that: „[T]his Directive lays down common rules on: (a) the disclosure of evidence on high-risk artificial intelligence (AI) systems to enable a claimant to substantiate a non-contractual fault-based civil law claim for damages; (b) the burden of proof in the case of non-contractual fault-based civil law claims brought before national courts for damages caused by an AI system”.

persons seeking compensation for damages caused by artificial intelligence systems. Other than that, the purpose of the AI Liability Directive is not to harmonize: „[g]eneral aspects of civil liability which are regulated in different ways by national civil liability rules, such as the definition of fault or causality, the different types of damage that give rise to claims for damages, the distribution of liability over multiple tortfeasors, contributory conduct, the calculation of damages or limitation periods”<sup>[42]</sup>. In the AI Liability Directive proposal, adopts the principle of minimum harmonization, allowing plaintiffs in cases of damage caused by AI systems to rely on more favorable rules of national law.

The AILD – similar as the PLD – contains facilitations in the area of disclosing evidence and legal presumptions supposed to support the claimant in the pursuit of compensation under the non-contractual liability regime for damages caused by AI systems. However, facilitations concerning disclosure of evidence have been provided only with regard to high-risk AI systems. The distinction between high-risk AI systems (and the definition of high-risk AI systems) and non-high-risk AI systems is taken from the AI Act. Under the AILD, both a plaintiff and a potential plaintiff are to be provided with facilitations with respect to evidentiary material<sup>[43]</sup>. The rights of a potential claimant were defined in two stages. Accordingly, in case of suspicion that a high-risk AI system caused damage, the potential claimant – prior to bringing the suit – will be in a position to request the AI system provider, a person subject to the obligations of a provider (under the AI Act) or the AI system’s user to disclose essential evidence in the disposal of such parties with regard to the specific AI system. If – despite a respective request from a potential claimant – such information is not provided, the potential claimant will be in a position to enforce the right to obtain information about an AI system in a court proceeding<sup>[44]</sup>. The potential plaintiff is obliged to submit sufficient facts and evidence to substantiate the claim for damages in the statement of claim submitted to the court. One should note the formulation: „[f]acts and evidence sufficient to support the plausibility of a claim for damages”. This is an indeterminate and undoubtedly evaluative expression. In this context, national courts adjudicating in non-contractual matters of civil law compensatory claims for damages

<sup>42</sup> The AILD proposal, point (10), 17. [https://commission.europa.eu/system/files/2022-09/1\\_1\\_197605\\_prop\\_dir\\_ai\\_en.pdf](https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf).

<sup>43</sup> See Art. 2(7) of the AILD proposal.

<sup>44</sup> See Art. 3(1) of the AILD proposal.

caused by AI systems will have a far-reaching decisive leeway in respect of which requests from potential plaintiffs deserve to be allowed. In case of the AILD – similarly as in the context of the PLD – it is clearly apparent that the EU legislator puts much emphasis on balancing the interests of potential claimants and defendants and on the adjustment of the level of legal protection afforded to both these groups to the – individually considered – factual circumstances of an individual case<sup>[45]</sup>. Analogous right will be available to the claimant already at the stage of judicial proceedings. Suppose the court orders the defendant to disclose or preserve evidence at the defendant’s disposal, and the defendant does not follow that order. In that case, the presumption of the defendant’s fault will apply<sup>[46]</sup>. In the AILD proposal, the EU legislator also provided for a rebuttable<sup>[47]</sup> legal presumption of a causal link between the defendant’s fault and the output produced by the AI system or the failure of the AI system to produce an output<sup>[48]</sup>. However, the presumption of a causal link will apply only where:

[a]ll of the following conditions are met: (a) the claimant has demonstrated or the court has presumed pursuant to Article 3(5), the fault of the defendant, or of a person for whose behaviour the defendant is responsible, consisting in the non-compliance with a duty of care laid down in Union or national law directly intended to protect against the damage that occurred; (b) it can be considered reasonably likely, based on the circumstances of the case, that the fault has influenced the output produced by the AI system or the failure of the AI system to produce an output; (c) the claimant has demonstrated that the output produced by the AI system or the failure of the AI system to produce an output gave rise to the damage<sup>[49]</sup>.

The Liability Directive proposal was limited to situations in which the requirements under the AI Act have not been fulfilled, as per the specific rules of Art. 4(2-3) of the AILD proposal. The possibility for the plaintiff to benefit from the legal presumption was strongly correlated to the safety requirements for high-risk AI systems contained in the AI Act. This means that if an AI system follows the safety requirements imposed on the system

<sup>45</sup> See Art. 3(3-4) of the AILD proposal.

<sup>46</sup> See Art. 3(5) of the AILD proposal.

<sup>47</sup> The defendant has the right to rebut the presumption laid down in Art. 4(1) of the AILD proposal. See Art. 4(7) of the AILD proposal.

<sup>48</sup> Art. 4(1) of the AILD proposal.

<sup>49</sup> Art. 4(1) of the AILD proposal.

according to the AI Act, the above-mentioned legal presumption will not apply. Accordingly, the European legislator clearly accentuates the high safety standards prescribed for high-risk AI systems under the AI Act. Furthermore:

[i]n the case of a claim for damages concerning a high-risk AI system, a national court shall not apply the presumption laid down in paragraph 1 where the defendant demonstrates that sufficient evidence and expertise is reasonably accessible for the claimant to prove the causal link<sup>[50]</sup>.

A situation in which the plaintiff can obtain easy access to the evidence and expert knowledge about a high-risk AI system will rather not be frequent in practice. In fact, the legal regime laid down in Art. 4(4) of the AILD proposal will provide an additional reinforcement of the legal situation of the defendant and, in turn, imply a weaker procedural position of the claimant. On the other hand, in relation to AI systems other than high-risk ones, the presumption laid down in Art. 4(1) of the AILD proposal, will apply only when the adjudicating court concludes that it is excessively difficult for the claimant to prove the causal link between the defendant's fault and the result produced by an AI system or the fact of not obtaining any result from such system<sup>[51]</sup>. Application of the presumption under Art. 4(1) of the AILD proposal will also be limited with regard to a defendant who has used the AI system exclusively as a part of his or her personal non-professional activity<sup>[52]</sup>. Therefore, according to the regulatory regime contained in the AILD proposal, the claimant still has to prove the defendant's fault<sup>[53]</sup> (or fault of a person for whose conduct the defendant is responsible) and the fact that the damage was caused by the AI system (i.e. that the damage was caused by a result generated by the AI system or by the fact of non-obtaining of any result from that system). On the other hand, the EU legislator offers a hand to the claimant by introducing a legal presumption of causal link between the defendant's fault and the damage caused by the AI system. However, application of that presumption by the adjudicating court is subject to many restrictions, which has been already

<sup>50</sup> Art. 4(4) of the AILD proposal.

<sup>51</sup> Art. 4(5) of the AILD proposal.

<sup>52</sup> Art. 4(6) of the AILD proposal

<sup>53</sup> Except for situations in which the court, under Art. 3(5) of the AILD proposal, has applied the presumption of fault.



discussed above. The liability of AI system operators has been included in the regulatory scope of the AILD, which clearly stipulates that such liability will continue to be based on fault.

## 5 | Reflections *de lege ferenda*

I join the voices of researchers who support the introduction of AI liability, understood as strict liability. For example, Christiane Wendehorst believes that the introduction of strict liability would be:

[a]n appropriate response to situations where significant and/or frequent harm may occur despite the absence of any fault, defect, malperformance or noncompliance. It may also be an appropriate response where such elements would be so difficult for the victim to prove that requiring such proof would lead to under-compensation or inefficiency. [...] The further extension of strict liability may be justified for AI applications because the «autonomy» and «opacity» of AI may give rise to exactly the kind of difficulties strict liability is designed to overcome<sup>[54]</sup>.

In my view, the solution to the problems related to liability for damage caused by AI systems would be achieved by departing from the liability regime for defective products in favour of a uniform, strict liability regime for damage caused by all AI systems. As a consequence, all producers of AI systems as well as professional operators of AI systems should incur risk-based liability. Just as previously, under Directive 85/374/EEC, a uniform liability regime was introduced for products understood as any movable items<sup>[55]</sup>, it is currently legitimate to introduce consolidated liability for damage caused by all and any AI systems. Thanks to such uniform liability principles, it would not be necessary to constantly update the list of high-risk AI systems. Moving on to details of such uniform, risk-based liability of both manufacturers and operators of all AI systems, in my opinion, there should be only two exonerating conditions: 1. force majeure; 2. exclusive

---

<sup>54</sup> Christiane Wendehorst, „Strict Liability for AI and other Emerging Technologies” *Journal of European Tort Law*, No. 2 (2020): 179.

<sup>55</sup> With the exception of primary agricultural products and game.

fault of the injured person or a third party for whom the injured person is responsible. As far as force majeure is concerned, it should be remembered that it will only be possible to invoke this circumstance in the case of external events, i.e. events that originate outside the AI system to whose operation the compensatory liability relates. If liability is to be exempt, the damage must be a consequence of the force majeure, which means that there is no adequate causal link between the operation of the system or device and the damage caused<sup>[56]</sup>. On the other hand, the exonerating conditions of the exclusive fault of the injured person or of a third person for whom the injured person is responsible will open a wide scope of exoneration for the defendants. This is the case since fault of the injured person (or of a third person) covers both intentional and unintentional fault. In this context, as far as unintentional fault is concerned, the defendant will be in a position to prove that the damage was caused because the injured person (or person for whom the injured person was responsible) did not acquire or did not maintain a level of knowledge necessary to duly operate the AI system, or that he or she acted under a misapprehension about the AI system's infallibility, as a result of which the outcomes generated by the system were not precisely verified by the injured person<sup>[57]</sup>. This will relate predominantly to such AI systems that could be used by a claimant and which the claimant used without involvement of the AI system's operator. As a result, the proposed liability model is not overly harsh on manufacturers and professional operators of AI systems. For appropriate legal protection of the injured parties, however, the essential part is that the defendant will be in a position to avoid liability only when the injured person's fault (or the fault of a third party for whom the injured party is responsible) is exclusive. That is when, apart from behavior – act or omission – of the injured person or a third party for whom the injured person is liable, there are no other factors contributing to the occurrence of damage or increasing its extent. On the other hand, in all those situations when the claimant (or third party for whom the claimant is responsible) only contributes to the occurrence of damage or to the increase of its scope, producers and operators of AI systems should incur compensatory liability, reduced according to the level of the injured person's (or the third party's) contribution to the occurrence the extent of the damage.

---

<sup>56</sup> Zbigniew Radwański, *Zobowiązania – część ogólna*, 2nd ed. (Warsaw: C.H. Beck, 1998), 77.

<sup>57</sup> Gerstner, „Liability Issues with Artificial Intelligence Software”, 249.

The burden of proving exonerating conditions should obviously be with the defendant. This is necessary to secure the position of the injured person. Moreover, the burden of proof should be reversed in relation to the causal link between the damage caused and operation of the AI system. Traditionally, in case of liability based on risk, the plaintiff bears the burden of proof with regard to the occurrence and extent of damage as well as the causal link between the damage and, for instance, service of a motor vehicle. However, in case of AI systems – having regard to their complexity, opacity and autonomy – it can be very difficult for the claimant to prove the causal relationship between operation of the system and the damage caused. This, in particular, will relate to those AI systems that are not incorporated into movable items. Thus, the burden of proof regarding the causal link between the operation of the AI system and the damage should be shifted from the plaintiff to the defendant. As a result, in order to avoid liability, the defendant would have to prove that there is no causal link between the operation of the AI system and the damage caused. It should be emphasized that the unified model of risk-based liability – as proposed above – would apply to all AI systems and to all their uses (whether direct or indirect). It would refer both to the use of AI systems through professional operators of those systems and to direct use, without involvement of any operators. The second type of use would naturally relate only to such AI systems that are intended for direct use.

In case of damage, it is the injured person that should decide whom he or she wishes to sue; the professional operator – if the injured person used the AI system through such operator – or directly the producer of the AI system (if such is the decision of the injured person or when there was no operator). Therefore, the injured person should have a choice. In all situations where there is an operator of an AI system, it will be easier for the injured person to identify the operator and seek compensation directly from the operator. In addition, the operator of an AI system should have the right to recourse compensation paid by the AI system's manufacturer in any case when the operator had paid compensation to the injured person. The principles of liability of a manufacturer of an AI system vis-à-vis the operator of that system should, however, be defined differently from that which can be found in the Regulation proposal. Notably, under Art. 12(3) of the Regulation proposal, it was envisaged that in situations when the operator of a defective AI system fully compensates for the damage to the injured person, the operator may seek recourse compensation from the manufacturer of that system under the terms of liability for a defective

product. As a consequence, liability of an AI system operator would be more severe than the liability of the AI system's manufacturer. In my opinion, the principles of liability of a manufacturer of an AI system and of the system's operator should be aligned. If a manufacturer wishes to avoid liability towards the operator of an AI system that compensated in full the damage suffered by the injured person, the manufacturer should prove that such damage was a consequence of force majeure or exclusive fault of the AI system's operator or a person for whom the operator is responsible<sup>[58]</sup>. If the manufacturer were unable to demonstrate any of the above-mentioned exonerating conditions, then the manufacturer would have to pay the compensation. The operator of a defective AI system should be relieved from the burden of proving the defect of the AI system that caused the damage and the causal link between the defect and the damage caused.

## 6 | Conclusion

Despite the high safety standards proposed in the AI Act, AI systems will undoubtedly cause damage, both in the material and immaterial dimension. This follows, among others, from the immanent characteristics of artificial intelligence: capacity to learn, to transform, to interact with humans, as well as its complexity, opacity, autonomy. The more complex a given AI system is, the more its properties merit from that complexity. Transformations that take place throughout the life cycle of an AI system will be a consequence of, among other things, interaction with the environment, including humans. In this context, AI systems that are safe upon being put into circulation – following the applicable legal requirements – have a potential to become dangerous later on in their lifecycle. This is the case since the potential to develop, represented by AI systems, is at the same time the potential to become a dangerous AI system. To face the current challenges, i.e. to make sure that the principles of liability for damages caused by AI systems create real – and not only illusory – protection of the

---

<sup>58</sup> The operator inappropriately handled the AI system (uploading to the AI system defective data or data other than of the highest quality; lack of ongoing updates; non-compliance with the specific instructions provided by the manufacturer; omission to maintain a due level of knowledge and/or skill necessary for the operation of the AI system), which gave rise to the damage.

injured parties and, at the same time, that they do not hinder the dynamic growth of artificial intelligence (safe AI), in my view, it is necessary to amend the EU legislation. Just as a couple of decades ago, in the European Union, separate terms of liability were introduced for defective products (Directive 85/374/EEC), presently we face a growing need to create a separate model of liability for damage caused by AI systems. The objective scope of such liability should cover all AI systems. On the other hand, as far as the subjective scope is concerned, in my opinion, all producers of AI systems and professional operators of AI systems (both back-end operators and front-end operators) should bear risk-based liability towards any person who has suffered injury or damage as a result of the physical or virtual operation, device or process controlled by an AI system.

The latest proposals for legislative amendments contained in the PLD proposal and the AILD proposal are rather conservative. With regard to civil liability for damage caused by AI systems, the EU legislator has not decided to introduce bold reforms, the nucleus of which, for example, could be seen in the European Parliament's resolution of October 20, 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence. According to the European Commission, it is still too early for any avant-garde reforms in the area of civil law liability of AI systems<sup>[59]</sup>. In consequence of the above, the EU legislator concentrated on such legal provisions that offer an answer to the problems with the liability for the discussed systems as reported so far. AI systems were unambiguously classified as products and covered by the civil liability regime for defective products (the Products Liability Directive proposal). The division has been preserved into the strict liability of manufacturers<sup>[60]</sup> of AI systems (liability for a defective product) and fault-based non-contractual liability attaching to providers and users of AI systems.

Although one can agree that the legal provisions contained in the PLD proposal and the AILD proposal are suited to the current development level of AI systems, it can be doubted whether the level of legal protection of persons injured by AI systems, as offered by the EU legislator, will be sufficient. This is the case since the European legislator strongly focused on balancing the interests of manufacturers, providers and users of AI systems with the interests of persons injured by such systems. The reason for such course of action was definitely the fear of slowing down the dynamic

---

<sup>59</sup> See the AILD proposal, Explanatory Memorandum, point 5, 14. [https://commission.europa.eu/system/files/2022-09/1\\_1\\_197605\\_prop\\_dir\\_ai\\_en.pdf](https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf).

<sup>60</sup> Manufacturers and other parties referred to in the PLD proposal.

development of the latest technologies and making the European Union a peripheral region in the sector of AI systems. Obviously, such fear is not groundless. However, the question arises as to whether the priority is the rapid development of AI systems and the maintenance of the European Union's competitiveness in relation to technological giants (e.g. China), or whether the most important task is to develop optimal legal conditions for the development of trustworthy and completely safe AI. As far as safety is concerned, however, the legal provisions contained in the AI Act inspire optimism. The solutions envisaged in that legislative act, at least in respect of high-risk AI systems, can really be considered ground-breaking and important in the perspective of the entire world. The AI Act sets completely new standards in the area of AI safety. Development of new technologies is important, and the newly created legal regulations are to be conducive in that regard. However, it is technology that should serve the human being and not the other way round. Therefore, the most important task is to protect human rights and European values. The intention behind putting people at the center is to develop safe and reliable AI<sup>[61]</sup>.

In summary, the proposals for legislative reform contained in the PLD proposal and the AI Liability Directive proposal are of a temporary nature and will certainly not be sufficient in the longer term. In any case, the European Commission is fully aware of this fact, as it emphasizes continuous monitoring of the situation in order

[t]o provide the Commission with information on incidents involving AI systems. The targeted review will assess whether additional measures would be needed, such as introducing a strict liability regime and/or mandatory insurance<sup>[62]</sup>.

---

<sup>61</sup> See: Ethics Guidelines for Trustworthy AI (2019), High-Level Expert Group on Artificial Intelligence. [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60419](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419); Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Empty „Building Trust in Human-Centric Artificial Intelligence”. COM(2019) 168 final. <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52019DC0168>; The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment (2020), High-Level Expert Group on Artificial Intelligence. <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>; European Parliament Resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). [https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html).

<sup>62</sup> The AILD proposal, Explanatory Memorandum, point 5, 14. [https://commission.europa.eu/system/files/2022-09/1\\_1\\_197605\\_prop\\_dir\\_ai\\_en.pdf](https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf).

It will only be possible to evaluate in retrospect and from a practical perspective to what extent and for how long the legal solutions contained in the PLD proposal and the AILD proposal prove efficient. For the time being, the legislative procedure in respect of the above-mentioned legislative acts has not yet been completed.

In this article, I have tried to show that the legislator should completely abandon liability for a defective product and instead adopt a uniform, risk-based liability regime for damage caused by all AI systems. Damage caused by an AI system should therefore be dissociated from a defect in that system. The risk-based liability regime should be limited to determining the damage and its extent, and whether there is a causal link between the operation of the AI system and the damage caused. On the other hand, it should be irrelevant whether the AI system was defective or not and, if so, what the nature of any defect was. The person injured by the operation of an AI system should also be compensated when the system was not defective – no defect in the technical and legal sense – or when it is impossible to establish what the possible defect consisted. In consequence, if the damage occurred in consequence of the operation of an AI system and there are no conditions exempting liability of the manufacturer or operator of such system (there is no exonerating condition), the injured party should receive compensation, regardless of whether the AI system was defective or not. Such a liability model, in combination with the rules on the safety of AI systems (e.g., AI Act), would allow to develop social confidence in the AI technology.

In my opinion, a uniform system of strict civil liability for damages caused by all and any AI systems<sup>[63]</sup>, risk-based and covering both manufacturers and professional operators of those systems is an adequate response to the challenges relating to the development of AI technologies. Critics of the position presented in this article could obviously raise an objection that risk-based liability would be too severe with regard to AI systems that do not pose a high risk. An additional objection could be made by saying that the introduction of risk-based liability for damages caused by AI systems would be favorable to the largest technological companies and, at the same time, detrimental to medium-sized and small firms or to start-up ventures. This relates to the opinion that only technological giants have the financial capacity to incur the costs of severe risk-based liability. The introduction of such a liability regime could therefore lead to a situation in which such

---

<sup>63</sup> This refers both to the most and to the least complex and autonomous AI systems.

giants monopolize the market for new technologies. In response to such objections, I would like to point out that, in my view, they are inaccurate. AI systems which do not pose a high risk do not bring such danger to health, life and property of natural persons as in the case of high-risk AI systems, correspondingly the damages caused by these systems will not be damage of great extent. Such damage will relate to individuals rather than entire communities. As a result, compensation to be paid by producers of AI systems not characterized by high risk will be proportionally lower. Only the frequency of their payment may be high, bearing in mind the increasing use of AI systems. However, it must be emphasized that severe risk-based liability will, in fact, encourage or even force manufacturers of AI systems to develop systems that are as safe as possible. This will relate to all manufacturers and all AI systems, starting from the least complex and the least autonomous to the most advanced high-risk AI systems. Since liability would be consolidated for manufacturers of all AI systems, all such manufacturers would also be affected by the pressure to minimize costs through developing increasingly safer AI. In this context, introduction of a uniform risk-based liability would have a very positive consequence: that is, a motivation to create ever safer AI systems. The manufacture of such systems, as well as payment of compensation for damage caused by AI systems, obviously implies a need to incur costs by those systems' producers. It should be remembered that manufacturers of AI systems verify their costs and ascertain the most cost-effective route: payment of compensation or investment in ever safer AI systems. Therefore, severe risk-based liability is conducive to raising safety standards. On the other hand, lenient liability for a defective product – which under the existing legal framework will not be enforceable in an increasing number of cases – categorically does not serve the same purpose. As regards the second objection signaled above, relating to the monopolization of the new technologies market by the largest firms operating in the area of artificial intelligence, it has been thoroughly examined by Anat Lior.<sup>[64]</sup> At this point, it seems legitimate to refer the reader to the scientific analysis carried out by that author. This analysis shows that large technological companies already enjoy a strong position in the AI industry. This is caused by many factors: the ability of such firms to attract the most talented specialists; having at their disposal huge data pools, gathered for years; having the greatest financial

---

<sup>64</sup> Anat Lior, „AI Strict Liability Vis-à-Vis AI Monopolization” *Columbia Science and Technology Law Review*, No. 1 (2020): 90-126.



resources. However, Lior shows that the introduction of risk-based liability will not change the status quo mentioned above to the detriment of small firms. On the other hand, the problem of monopolizing the market should be solved by appropriate regulations specifically devoted thereto. However, abandonment of risk-based liability, most definitely, cannot be regarded as one of such regulations. There is no proof that by abandoning a risk-based liability regime for AI systems the legislator would improve the situation of small firms trying their hand in the AI sector.<sup>[65]</sup> Lior's argument leaves no doubt that we should not abandon the risk-based liability model for damages caused by AI systems.

## Bibliography

- Fairgrieve Duncan, Geraint Howells, Peter Møgelvang-Hansen, Gert Straetmans, Dimitri Verhoeven, Piotr Machnikowski, André Janssen, Reiner Schulze, „Product Liability Directive”, [in:] *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, ed. Piotr Machnikowski. 46. Cambridge, Antwerp, Portland: Intersentia, 2016.
- Gerstner Maruerite E., „Liability Issues with Artificial Intelligence Software” *Santa Clara Law Review*, No. 1 (1993): 239-269. <https://digitalcommons.law.scu.edu/lawreview/vol33/iss1/7/>.
- Gilead Israel, „Product Liability in Israel”, [in:] *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, ed. Piotr Machnikowski. 546-547. Cambridge, Antwerp, Portland: Intersentia, 2016.
- Jagielska Monika, „Odpowiedzialność za sztuczną inteligencję [Liability for Artificial Intelligence]”, [in:] *Prawo sztucznej inteligencji [The Law of Artificial Intelligence]*, ed. Luigi Lai, Marek Świerczyński. 69-79. Warsaw: C.H. Beck, 2020.
- Jagielska Monika, *Odpowiedzialność za produkt [Product Liability]*. Warsaw: Wolters Kluwer, 2009.
- Lior Anat, „AI Strict Liability Vis-à-Vis AI Monopolization” *Columbia Science and Technology Law Review*, No. 1 (2020): 90-126. <https://doi.org/10.52214/stlr.v22i1.8055>.
- European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, ed. Piotr Machnikowski. Cambridge, Antwerp, Portland: Intersentia, 2016.

<sup>65</sup> Ibidem, 125-126.

- Masci Pietro, „The History of Insurance: Risk, Uncertainty and Entrepreneurship” *Journal of the Washington Institute of China Studies* No. 3 (2011): 25-68. [https://www.academia.edu/45474658/The\\_History\\_of\\_Insurance\\_Risk\\_Uncertainty\\_and\\_Entrepreneurship](https://www.academia.edu/45474658/The_History_of_Insurance_Risk_Uncertainty_and_Entrepreneurship).
- Radwański Zbigniew, *Zobowiązania – część ogólna [Obligations – General]*, 2nd ed. Warsaw: C.H. Beck, 1998.
- Rojszczak Marcin, Marcin Rojszczak, „Prawne aspekty systemów sztucznej inteligencji. Zarys problemu [Legal Aspects of Artificial Intelligence Systems: Outline of the Problem]”, [in:] *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane [Artificial Intelligence, Blockchain, Cybersecurity and Personal Data: Selected Issues]*, ed. Kinga Flaga-Gieruszyńska, Jacek Gołaczyński, Dariusz Szostek. 1-22. Warsaw: C.H. Beck, 2019.
- Russell Stuart, *Human Compatible: AI and the Problem of Control*. UK, USA, Canada, Ireland, Australia, India, New Zealand, South Africa: Penguin Books, 2019.
- Smuha Nathalie A., Emma Ahmed-Rengers, Adam Harkens, Wenlong Li, James MacLaren, Riccardo Piselli, Karen Yeung. *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission’s Proposal for an Artificial Intelligence Act*. Birmingham: University of Birmingham, 2021. <http://dx.doi.org/10.2139/ssrn.3899991>.
- Swanson Greg, „Non-Autonomous Artificial Intelligence Programs and Products Liability: How New AI Products Challenge Existing Liability Models and Pose New Financial Burdens” *Seattle University Law Review*, No. 3 (2019): 1201-1222. <https://digitalcommons.law.seattleu.edu/sulr/vol42/iss3/11/>.
- Sykes Alan O., „The Law and Economics of «Forced» Technology Transfer and Its Implications for Trade and Investment Policy (and the U.S.–China Trade War)” *Journal of Legal Analysis*, No. 1 (2021): 127-171. <https://doi.org/10.1093/jla/laaa007>.
- Veale Michael, Frederik Zuiderveen Borgesius, „Demystifying the Draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach” *Computer Law Review International* No. 4 (2021): 97-112. <https://doi.org/10.9785/crl-2021-220402>.
- Vladeck David C., „Machines without Principals: Liability Rules and Artificial Intelligence” *Washington Law Review*, No. 1 (2014): 117-150. <https://digitalcommons.law.uw.edu/wlr/vol89/iss1/6>.
- Wendehorst Christiane, „Strict Liability for AI and other Emerging Technologies” *Journal of European Tort Law*, No. 2 (2020): 179. <https://doi.org/10.1515/jetl-2020-0140>.

