

Universidades Lusíada

González, José A.R.L., 1965-
González, Maria Margarida L.
González, José Martinho

Responsabilidade civil pela utilização de dispositivos médicos dotados de inteligência artificial (IA)

<http://hdl.handle.net/11067/8355>
<https://doi.org/10.34628/WFHE-TH76>

Metadados

Data de Publicação

2025

Resumo

Existem já diversos dispositivos dotados de IA capazes de desempenhar um papel significativo na melhoria da qualidade da assistência médica em diferentes áreas, desde o diagnóstico até à execução de complexas cirurgias. A sua utilização suscita, no entanto, sérias interrogações, especialmente quando dela resultem danos para os pacientes. Este artigo tem em vista, sobretudo, proceder ao respetivo levantamento e à análise crítica das soluções que habitualmente têm sido propostas....

There are already several AI-enabled devices that can play a significant role in improving the quality of healthcare in different areas, from diagnosis to complex surgeries. However, its use raises serious questions, especially when it results in harm to patients. This article aims, above all, to carry out a survey and critical analysis of the solutions that have usually been proposed....

Editor

Universidade Lusíada Editora

Palavras Chave

Responsabilidade (Direito), Instrumentos e aparelhos médicos, Inteligência Artificial

Tipo

article

Revisão de Pares

yes

Coleções

[ULL-FD] LD, s. 2, n. 34 (2025)

Esta página foi gerada automaticamente em 2026-10-30T04:54:51Z com informação proveniente do Repositório

RESPONSABILIDADE CIVIL PELA UTILIZAÇÃO DE DISPOSITIVOS MÉDICOS DOTADOS DE INTELIGÊNCIA ARTIFICIAL (IA) ¹

CIVIL LIABILITY FOR THE USE OF MEDICAL DEVICES EQUIPPED WITH ARTIFICIAL INTELLIGENCE (AI)

José Lorenzo González ²

Maria Margarida L. González ³

José Martinho González ⁴

DOI: <https://doi.org/10.34628/WFHE-TH76>

Resumo: Existem já diversos dispositivos dotados de IA capazes de desempenhar um papel significativo na melhoria da qualidade da assistência médica em diferentes áreas, desde o diagnóstico até à execução de complexas cirurgias. A sua utilização suscita, no entanto, sérias interrogações, especialmente quando dela resultem danos para os pacientes. Este artigo tem em vista, sobretudo, proceder ao respetivo levantamento e à análise crítica das soluções que habitualmente têm sido propostas.

Palavras-chave: Responsabilidade civil; Inteligência artificial; Responsabilidade médica.

Abstract: There are already several AI-enabled devices that can play a significant role in improving the quality of healthcare in different areas, from diagnosis to complex surgeries. However, its use raises serious questions, especially when it results in harm to patients. This article aims, above all, to carry out a survey and critical analysis of the solutions that have usually been proposed.

¹ Publicado na Revista do Ministério Público, n.º 183 (Jul-Set) 2005, Ano 46.

² Professor Catedrático da Universidade Lusíada – Lisboa Investigador integrado do CEJEIA

³ Médica Especialista em Ginecologia e Obstetrícia (Hospital de Cascais)

⁴ Mestre em Direito; Advogado-Estagiário (PLMJ)

Keywords: Civil liability; Artificial intelligence; Medical liability.

§ 1. A inteligência artificial (IA) está a introduzir-se rapidamente em todos as áreas da vida em que a tecnologia pode irromper e também, portanto, na prática médica ⁵. Seja para previsão de risco, para diagnóstico ou para recomendação de tratamentos ⁶.

⁵ “In health and life sciences the first AI developments were in the 1960s and 1970s, with a system called DENDRAL designed to assist chemists in identifying structures of organic molecules. DENDRAL did not apply specifically to health, but it led to other systems with more specific healthcare applications, such as MYCIN. This was a computer-based consultation system used by doctors for the diagnosis and therapy of patients with bacterial infections. There was MOLGEN, designed to assist scientists in developing complex experiment plans and analysing molecules in the domain of molecular biology. Another system was INTERNIST-I, an experimental program capable of making multiple and complex diagnoses in internal medicine” (Barry Solaiman and I. Glenn Cohen, *A framework for health, AI and the law*, Research Handbook on Health, AI and the Law, Elgar Publishing, 2024, pág. 2).

⁶ Perguntado o ChatGPT (versão base), em 30/03/2025, sobre “medical decision and AI”, obteve-se a seguinte resposta: “AI has increasingly been used in medical decision-making, offering new tools for healthcare professionals. Here are some key ways AI is being integrated into medical decisions:

1. Diagnostics

Image recognition: AI algorithms can analyze medical images (e.g., X-rays, MRIs, CT scans) with high accuracy. These tools can help detect issues like tumors, fractures, or diseases such as pneumonia and breast cancer.

Pattern recognition: AI models can identify patterns in data that may not be immediately obvious to doctors, enabling earlier diagnosis of conditions like heart disease or diabetes.

2. Treatment Planning

Personalized treatment: AI can help analyze a patient’s medical history, genetic information, and other factors to recommend personalized treatment plans, ensuring a more tailored approach to care.

Predicting treatment outcomes: By analyzing large datasets, AI can help predict how well certain treatments may work for specific patients, allowing doctors to make better-informed decisions.

3. Drug Discovery and Development

AI is helping researchers identify new drug compounds faster. Algorithms can analyze vast amounts of data on chemical structures and biological effects, speeding up the discovery of potential treatments.

4. Clinical Decision Support Systems (CDSS)

AI-based systems provide healthcare professionals with evidence-based recommendations. These systems can assist in diagnosing conditions, suggesting treatments, and even identifying potential risks, all while ensuring they have the latest research data available.

5. Predictive Analytics

AI models can predict the likelihood of a patient developing certain conditions based on factors

Alguns dispositivos de IA médica existem já com aptidão para tomar decisões de forma totalmente independente, sem intervenção humana. Um dispositivo médico dotado de IA é capaz, por exemplo, de detetar a retinopatia diabética, excluindo, com vantagem, a necessidade de obter confirmação, dada por um especialista humano, sobre a interpretação de certa imagem ⁷; ou é igualmente capaz de detetar, com grande antecipação, o provável desenvolvimento de cancro do pulmão ⁸. E, no futuro próximo, apenas é de esperar que muito mais se venha a fazer de modo autónomo. Por necessidade, sobretudo. A carência de médicos, a elevação dos *standards* a que a prestação de cuidados de saúde se submete, as distâncias ou as localizações geográficas em que os pacientes se encontram, os custos de construção e manutenção de unidades hospitalares ⁹, assim o determinarão.

like lifestyle, genetics, and medical history. This can help with early intervention and preventative care.

6. Robotic Surgery

AI is integrated into robotic systems used in surgeries, allowing for more precise movements, faster recovery times, and reduced human error.

7. Natural Language Processing (NLP)

AI systems using NLP can process and analyze clinical notes, research papers, and patient records to extract valuable information for decision-making. This can help in recognizing symptoms, treatment options, and drug interactions”.

⁷ “EyeDiagnosis’s IDx-DR software for diabetic retinopathy is an example of leveling-up that AI can bring to medical diagnosis. (...) The current standard of care is for patients with diabetes to visit an ophthalmologist yearly to check for signs of retinopathy, so that treatment can begin before the retina worsens. But this requires regularly visiting an ophthalmologist, which is not easy or even possible for many patients. EyeDiagnosis has developed a system that enables primary care physicians (or other non-specialist practitioners) to use an essentially automated camera to take images of the retina; those images are then analyzed by a machine-learning algorithm trained on a gold-standard dataset of retina images (annotated by expert ophthalmologists). The algorithm returns a diagnosis of more-than-mild diabetic retinopathy, in which case the patient should seek further care, or not, in which case the patient should ideally be retested in a year” (W. Nicholson Price, Medical AI and Contextual Bias, Harvard Journal of Law & Technology, Volume 33, Number 1, 2019, pág. 75).

⁸ Cf. v.g. Peter G. Mikhael – Jeremy Wohlwend – Adam Yala – Ludvig Karstens – Justin Xiang – Angelo K. Takigami – Patrick P. Bourgouin – PuiYee Chan – Sofiane Mrah – Wael Amayri – Yu-Hsiang Juan – Cheng-Ta Yang – Yung-Liang Wan – Gigin Lin – Lecia V. Sequist – Florian J. Fintelmann – Regina Barzilay, Sybil: A Validated Deep Learning Model to Predict Future Lung Cancer Risk From a Single Low-Dose Chest Computed Tomography, Journal of Clinical Oncology, 2023, Jan 12, 41 (12), 2191/2200.

⁹ “A key promise of medical AI is its ability to democratize medical expertise, allowing providers of all sorts to give care that otherwise might be beyond their capacity” (W. Nicholson

§ 2. Em geral, o algoritmo de IA pode apresentar-se como “bloqueado” ou como “adaptável”. Os dispositivos médicos que utilizam algoritmos “bloqueados” produzirão o mesmo resultado sempre que lhe seja aplicada a mesma “entrada”, não sendo o seu uso recorrente suscetível de lhes provocar alterações procedimentais. Ao invés, os sistemas “adaptáveis”, que aprendem e se ajustam continuamente, são passíveis de fornecer resultados diferentes dos antecipados no momento da sua aprovação regulamentar¹⁰. São estes últimos suscetíveis, por isso, de fazer despontar o chamado *black box effect*¹¹ e trazer a necessidade de *explicabilidade*.

O que antecede suscita de imediato, porém, uma óbvia questão: quem responsabilizar quando algum procedimento corra deficientemente e daí advenham danos? Ou seja: quando um paciente seja de alguma forma lesado – devido *v.g.* a errado diagnóstico ou tratamento – e nesse resultado a IA se encontrar envolvida, seja a que título for, quem deverá obrigar-se à indemnização e com que fundamento?

Para já, a hipótese mais realista de utilização da IA na medicina será no âmbito do chamado “apoio à decisão clínica” (*clinical decision support* –

Price, *Medical AI and Contextual Bias*, Harvard Journal of Law & Technology, Volume 33, Number 1, 2019, pág. 66).

¹⁰ “An AI system utilizing machine learning employs an algorithm programmed to learn («learner algorithm») from data referred to as «training data». The learner algorithm will then automatically adjust the machine learning model based on the training data. A «continuous learning system» updates the model without human oversight as new data is presented, whereas «locked learners» will not automatically update the model with new data. In health care, it is important to know whether the learner algorithm is eventually locked or whether the learner algorithm continues to learn once deployed into clinical practice in order to assess the systems for quality, safety, and bias. Being able to trace the source of training data is critical to understanding the risk associated with applying a health care AI system to individuals whose personal characteristics are significantly different than those in the training data set” (WMA Statement on Augmented Intelligence in Medical Care – Adopted by the 70th WMA General Assembly, Tbilisi, Georgia, October 2019, Preamble).

¹¹ “For decades, physicians have used clinical guidelines and decision support systems based on statistical models as a tool in the practice of medicine. While early statistical models curated medical knowledge to create static clinical decision support systems, modern AI/ML builds highly complex programs involving extensive data, designed to learn from previous datasets, decipher new patterns, and relies on adaptive models that may change recommendations over time based on inputs. The machine’s vast layers of neural networks are often referred to as a «black box», or opaque based on the lack of transparency and explainability to outside observers” (Katherine Drabiak, *Leveraging law and ethics to promote safe and reliable AI/ML in healthcare*, Frontiers in Nuclear Medicine, 2022, pág. 3).

CDS ¹²). Nele incluir-se-ão, todavia, tanto funções relativamente elementares – v.g. alertas e lembretes para prestadores e/ou doentes – como funções de alta complexidade – v.g. o diagnóstico de pacientes ¹³.

Ante o recurso a estes sistemas de *clinical decision support*, o médico confrontará, no final, o resultado que o algoritmo lhe fornecer com os seus próprios conhecimentos. O que significa que lhe cabe assumir a responsabilidade pela decisão diagnóstica final. Mas é claramente antecipável que, principalmente por necessidade ou urgência (como antes se dizia), este estado de coisas não permaneça assim por muito tempo, em especial quando a IA médica passe a beneficiar de uma confiabilidade geral. Aí, contudo, havendo danos e lesados, será imprescindível definir a quem imputar a obrigação de indemnizar quando o procedimento decorra impreteritamente.

§ 3. A forma como esta responsabilização se construir decerto influenciará o comportamento dos prestadores de serviços médicos sempre que lhes couber ajuizar sobre a adoção de recomendações fornecidas pela IA; interferirá nas decisões de gestores hospitalares relativas à implementação de ferramentas de IA para uso médico; assim como, antes do mais, determinará o comportamento dos fabricantes de tais equipamentos.

Se ela se encontrar convenientemente desenvolvida, o recurso à IA pode facilitar a colocação da decisão clínica próxima do padrão de cuidado exigível pelas *legis artis*. Eventualmente, até, superá-lo. Em tal

¹² “Clinical decision support (CDS) is broadly defined as the provision of intelligent assistance to clinicians, medical staff, and patients. CDS can include low-level functions like computerized alerts and reminders for providers and patients, or high-level functions like patient diagnosis. Norman et al. describe a dual process of diagnostic reasoning, where physicians engage in (1) a non-analytic or unconscious process of hypothesis generation, and/or (2) a conscious, analytic process of hypothesis testing. The latter is an extensive computational process, and has motivated efforts to develop AI-based CDS systems for diagnostic support.

In such support systems, a physician cross-checks the algorithmic output against their internal knowledge, but takes responsibility for the final diagnostic decision. Our work takes a similar approach of augmenting, instead of automating, the job of physicians” (Mike Schaeckermann, Graeme Beaton, Elaheh Sanoubari, Andrew Lim, Kate Larson, Edith Law, *Ambiguity-aware AI Assistants for Medical Data Analysis*, CHI 2020 Paper CHI 2020, April 25–30, 2020, Honolulu, HI, USA, pág. 2).

¹³ O diagnóstico compreenderá, em geral, duas etapas: (i) um processo, em maior ou menor medida, inconsciente de geração de hipóteses; (ii) e, de seguida, uma abordagem analítica consciente destinada a testar as hipóteses antes formuladas. A última é, genericamente, um processo computacional suscetível de alicerçar-se em IA.

condição, adotar os conselhos fornecidos pela IA reduzirá o risco de responsabilização. E certamente chegará o momento em que, ao invés, o médico se sujeitará a responsabilidade quando se desvie das recomendações padronizadas de tratamento geradas pela IA ¹⁴. Com efeito, assim que o tratamento médico assistido por IA atingir maior precisão do que os tratamentos efetuados sem o seu envolvimento, o respetivo uso constituirá o (novo) padrão de cuidado a observar na assistência médica, tornando os prestadores de cuidados de saúde, ao invés, responsáveis pela não utilização de IA ¹⁵.

§ 4. Recentes desenvolvimentos no domínio da IA aplicada aos cuidados de saúde prometem resolver muitos dos atuais problemas globais em matéria de promoção da saúde e de gestão de desafios gerais nesse domínio. Tal contem, todavia, implicações éticas, jurídicas e sociais que merecem cuidado e atenção, tanto na sua implementação, como na sua execução.

Numa declaração da 70.^a Assembleia Geral, em 2019, a Associação Médica Mundial também reconheceu o vasto potencial da IA. Sugeriu, no entanto, a utilização da locução “inteligência aumentada” como uma forma limitada de IA, assim sublinhando o seu (simples) papel de *apoio* à decisão humana ¹⁶. Esta variação de terminologia descreve com precisão o presente objetivo de fazer coexistir a IA com a tomada de decisões humanas, apoiando e melhorando a construção de juízos clínicos.

¹⁴ “The standard of care is especially protective of physicians because it is not unitary; if physicians can show that they have followed the standard as practiced by a ‘«respectable minority»’ of the medical field, that too will be enough to prevent liability. So even when the use of AI becomes the standard of care, older rubrics will protect physicians who follow them for some time” (W. Nicholson Price – Sara Gerke – I. Glenn Cohen, *How Much Can Potential Jurors Tell Us About Liability for Medical Artificial Intelligence?*, The Journal of Nuclear Medicine, Vol. 62, No. 1, January 2021, pág. 16).

¹⁵ “Big data and machine learning can help health-care providers explore new biological relationships and new methods of treatment, automate many low-level tasks that fill providers’ days, and raise the general level of care by allowing many types of providers to access expertise through the intermediary of medical AI” (W. Nicholson Price, *Medical AI and Contextual Bias*, Harvard Journal of Law & Technology, Volume 33, Number 1, 2019, pág. 70).

¹⁶ “In health care, a more appropriate term is «augmented intelligence», an alternative conceptualization that more accurately reflects the purpose of such systems because they are intended to coexist with human decision-making” (WMA Statement on Augmented Intelligence in Medical Care – Adopted by the 70th WMA General Assembly, Tbilisi, Georgia, October 2019, Preamble).

Em distinta direção, mas revelando idêntica preocupação, também o World Economic Forum definiu as chamadas “four key principles in our Ethical use of Data & Technology Policy”¹⁷:

1. *Respect humanity*: Respect humanity and maintain trust with society through deploying and using data and technology in ways that promote diversity and inclusion, respect human rights and benefit patients and society.

2. *Be transparent and collect fairly*: Describe clearly and simply why and how the organization collects data and what it does with data and technology.

3. *Use responsibly*: Use data and technology responsibly and be accountable for the appropriate management and protection of them in accordance with your scope of responsibility.

4. *Protect data and technology*: Apply risk-based security to protect and ensure the confidentiality, integrity and availability of data and technology throughout their lifecycle. Prevent data loss”.

§ 5. Embora a IA possa trazer muitos benefícios para os cuidados de saúde¹⁸, não consegue, para já, substituir componentes críticos da relação médico-doente, cujos princípios fundamentais são a confiança, o conhecimento, a consideração, a lealdade e a empatia. Médicos empáticos e emocionalmente empenhados, ao estimular confiança, obtêm maior eficácia terapêutica e diagnósticos mais completos, diminuindo a ansiedade do doente e promovendo o respetivo envolvimento na sua própria saúde. Será importante preservar este relacionamento como um valor central em todo o desenvolvimento da IA médica¹⁹. A prática da medicina é caracterizada por uma abundância de qualidades intangíveis em que o médico

¹⁷ https://www.weforum.org/stories/2025/01/the-4-principles-of-responsible-ai-in-medicine/?gad_source=1&gclid=Cj0KCQjw2ZfABhDBARIsAHFTxGw3WV6HtRLttrt-op3JSJhlsYpkZLi3HugNXXWUO-tyszLhFkbUUYgaAsnFEALw_wcB

¹⁸ “Complexity and diversity are hard for humans. The human mind can only balance several factors – perhaps two to five – simultaneously. There is an inherent mismatch between the mathematical optimization in high-dimensionality characteristic of machine learning and the [more limited] demands of human scale reasoning. Machine brains can weigh thousands of factors at once and cope with diversity and nuance unemotionally” (Barbara J. Evans, *Rules for robots, and why medical AI breaks them*, Journal of Law and the Biosciences, 1–35, 2023, págs. 2/3).

¹⁹ “Health care AI systems do not replace the need for the patient-physician relationship. Such systems augment physician-provided medical care and do not replace it” (WMA Statement on Augmented Intelligence in Medical Care – Adopted by the 70th WMA General Assembly, Tbilisi, Georgia, October 2019, Preamble).

humano sempre se destacará em relação às máquinas. Isto inclui empatia, compaixão, toque, contacto visual e captação de sinais sociais diferenciados (como *v.g.* o tom de voz). Ainda que não seja impossível dotar máquinas providas de IA com algumas destas qualidades, o ser humano há-de, normalmente, querer ser cuidado por outro ser humano. E a interação humana, especialmente na relação médico-paciente, é vital. Além disso, o médico humano tem uma abordagem não linear na execução do seu trabalho capaz de incutir, naquela relação, níveis de personalização não permitidos por uma abordagem unicamente orientada por dados. As doenças podem ter características semelhantes suscetíveis de fácil reconhecimento pela IA, mas os doentes e os seus estilos de vida são distintos. Os diagnósticos e tratamentos exigem, por isso, competências que as máquinas, para já, não possuem.

Indiscutivelmente, a IA parece apresentar consideráveis vantagens na melhoria da precisão do diagnóstico, no tratamento de doenças e na gestão clínica e hospitalar, impulsionando avanços na investigação em saúde e nas inovações farmacêuticas, bem como na otimização da gestão dos serviços e recursos de saúde. Mas, claro, também exhibe contraindicações. Normalmente, os algoritmos de IA são validados mediante análises retrospectivas executadas sobre o conjunto dos dados arquivados. O objetivo é avaliar a sua capacidade *preditiva* para uma tarefa médica específica (como *v.g.* decidir qual o melhor tratamento para um doente e antecipar possíveis resultados de tratamentos anteriores). Isso não se afigura suficiente, contudo, quando se considere: (i) a sua precisão em relação a dados não revistos; (ii) a complexidade da tomada de decisão na prática clínica.

§ 6. Há problemas gerais (isto é, característicos da IA) a resolver.

Primeiro, a chamada (e já referida) *explicabilidade*. Dada a sua complexidade, os sistemas de IA são normalmente opacos no que respeita ao respetivo modo de funcionamento. Ora, em áreas onde a IA seja responsável por decisões importantes, é imperativo que o sistema automatizado de tomada de decisão seja interpretável. O campo de XAI [*explainable artificial intelligence*] surgiu precisamente como resposta a esta procura de transparência. Em medicina, isso significará, por exemplo, que os algoritmos sejam capazes de proceder à demonstração do raciocínio subjacente à tomada de decisão, de diagnóstico e de tratamento, relativa a uma doença que haja sido detetada. Será necessário, para isso, que os sistemas de IA tenham aptidão para produzir explicações em linguagem natural semelhantes às humanas, emulando-as.

Segundo, a *confiança*. Ou melhor, a sua falta, decorrente do chamado enviesamento algorítmico é, sem dúvida, a barreira mais significativa

para a adoção de sistemas de IA destinados a apoiar a decisão médica. Os especialistas humanos também podem falhar, mas parece, para já, que o problema será mais agudo ante, em geral, sistemas artificiais ²⁰.

Terceiro, a *igualdade* no acesso aos cuidados de saúde. Para já, só hospitais dotados de grandes recursos financeiros ou centros de investigação de ponta permitirão aos seus pacientes o apelo a dispositivos médicos dotados de IA. Enquanto não se tornarem de *open access*, os respetivos fabricantes ou *developers* apenas estarão disponíveis para os ceder a terceiros mediante avultadas compensações pecuniárias. O que oferece o perigo, portanto, de a IA médica servir (ao menos, de início) para amplificar, ainda mais, as disparidades no acesso aos cuidados de saúde ²¹.

§ 7. Reconhecer o potencial de erros induzidos pela IA é uma hipótese a dever encarar-se realisticamente e, em tais cenários, a questão de saber onde reside a responsabilidade ²² torna-se particularmente relevante.

²⁰ “There have been numerous examples of AI failures or reported bias primarily related to social norms such as gender and race. The field of biomedicine has not been immune to these biases. For instance, an ML algorithm that was developed for predicting the risk of pneumonia, it counterintuitively suggested that patients with pneumonia and asthma will be at a lower risk of death than patients with pneumonia but without asthma. Similar controversial examples were also noted in the case of skin cancer risk prediction, where the presence of a ruler in the image may have been a cue for the ML algorithm of high risk, or in the case lung disease where the appearance of a tube in a chest X-ray was an indication of disease severity” (Katherine Drabiak – Skylar Kyzer – Valerie Nemo – Issam El Naqa, *AI and machine learning ethics, law, diversity, and global impact*, Br J Radiol, 2023, pág. 2).

²¹ “Democratization of the expertise needed to recommend treatments, though, does not always translate to democratization of the expertise needed to actually treat. If medical AI recommends taking tablet A over tablet B for a particular patient, as long as both tablets are available, that recommendation may be easy to follow. But if medical AI recommends a complicated surgery, successful implementation depends on the presence of a skilled surgeon” (W. Nicholson Price, *Medical AI and Contextual Bias*, Harvard Journal of Law & Technology, Volume 33, Number 1, 2019, pág. 77).

²² “A recurrent problem in adapting law to artificial intelligence programs is how the law should regulate the use of entities that lack intentions. Many areas of the law, including freedom of speech, copyright, and criminal law, make liability turn on whether the actor who causes harm (or creates a risk of harm) has a certain intention or *mens rea*. But AI agents – at least the ones we currently have – do not have intentions in the way that humans do. If liability turns on intention, that might immunize the use of AI programs from liability” (Ian Ayres – Jack M. Balkin, *The Law of AI is the Law of Risky Agents without Intentions*, University of Chicago Law Review Online, 2024, pág. 1).

Os sistemas de IA de última geração são capazes de fornecer dados quantitativos contendo estimativas de certeza/incerteza (v.g. 80% de probabilidade que o estado de saúde de um paciente seja “anómalo”). Estas estimativas derivam normalmente de distribuições de probabilidades executadas *a posteriori*. No entanto, elas nem sempre antecipam as discordâncias de maneira fiável. Além disso, as representações numéricas da incerteza realizadas através de IA, só por si, podem não bastar para que aos especialistas humanos os respetivos motivos adquiram sentido. E o desacordo (entre especialistas) na análise de dados médicos pode assentar em diversos fundamentos: definição ambígua dos problemas, pluralidade de respostas corretas, enviesamento do raciocínio, etc.

Aplicando as regras relativas à responsabilidade do produtor, os pacientes que forem prejudicados por produtos médico-cirúrgicos que não sejam “razoavelmente seguros”, devido a falhas de conceção ou a deficiente fabrico, têm o direito de obter a competente indemnização ou compensação. O produtor responderá pelas eventuais deficiências – e, portanto, pelos danos daí decorrentes – patenteadas por medicamentos (sujeitos a prescrição médica) ou por dispositivos médicos. Reconhece-se, no fundo, que a utilização destes produtos envolve riscos inevitáveis, necessitando até, por isso, de autorização prévia como pré-requisito para a sua introdução no mercado. Mas será esta a solução preferível?

A pergunta transmuta-se, assim, na seguinte: “Quem será responsável por um erro de diagnóstico ou de tratamento imputável a uma tecnologia de IA defeituosa?”²³.

No estado atual de conhecimentos, a IA tem exibido um amplo espectro de desempenho, passando, sem aviso prévio, de capacidades cognitivas altamente avançadas para um comportamento marcadamente básico ou até rudimentar²⁴. Independentemente dos esforços para atenuar

²³ “Responsibility within jurisprudence is classified into three main categories: (1) Accountability, (2) Culpability, and (3) Liability. The term «accountability» refers to the capability of a system to provide an explanation for its actions, encompassing a detailed account of its input(s), internal processes, and output(s). A person is deemed «culpable» for a committed punishable action if this action was either intentionally and freely undertaken, or if the individual failed to take necessary measures to prevent the act. Given these definitions, an AI program should be «accountable» but cannot be considered «culpable» for its actions, as it lacks free will, consciousness, and the capacity for autonomous decision-making” (Contaldo, M.T. – Pasceri, G. – Vignati, G. – Bracchi, L. – Triggiani, S. – Carrafiello, G, *AI in Radiology: Navigating Medical Responsibility*, Diagnostics, 2024, 14, 1506, pág. 5).

²⁴ “For instance, an ML algorithm that was developed for predicting the risk of pneumonia, it counterintuitively suggested that patients with pneumonia and asthma will

as “tendências” da IA, persistem limitações intrínsecas em todas as suas construções. A atribuição de responsabilidades aos respetivos criadores ou operadores é potencialmente comprometida pela autonomia dos sistemas de IA, que podem funcionar de forma independente e aprender (ou, pelo menos, desenvolver), por si, novos padrões de comportamento. A implantação de tecnologias de IA é capaz de provocar, portanto, um vazio de responsabilização, sem um agente identificável.

A segurança dos produtos é, contudo, uma componente chave da estrutura jurídica a criar para a IA médica ²⁵. Tornar-se-á necessária, por isso, uma avaliação prévia antes de ocorrer a sua colocação em mercado, de modo a garantir a respetiva conformidade com as normas elementares de segurança. Estas, em geral, têm em vista minimizar o risco de danos provocados por um produto defeituoso, visando oferecer um elevado nível de confiança para os utilizadores. A maioria das aplicações de IA médica qualificam-se como *software* destinado pelo fabricante a ser utilizado, para fins médicos específicos, sobre seres humanos e, portanto, como dispositivos médicos (ou como partes componentes de um dispositivo médico). Com base na sua classificação de risco, tornam-se indispensáveis, assim, procedimentos de avaliação da conformidade (certificação, revisão, evidência clínica para comprovação da precisão e da fiabilidade, etc.) ²⁶.

§ 8. O reconhecimento dos direitos fundamentais tem em vista, não só tutelar os cidadãos contra a intromissão do Estado na sua esfera jurídica, como também obrigar o Estado a conferir proteção efetiva contra o exercício intrusivo de certas liberdades alheias. O Estado cumpre estas chamadas “obrigações de proteção” através, por exemplo, da promulgação de legislação apropriada que se aplique às relações entre sujeitos de direito privado ou criando procedimentos de aprovação específicos para colocar no mercado bens ou serviços que possam pôr em perigo os direitos fundamentais dos seus utilizadores. É por isso que as “obrigações de

be at a lower risk of death than patients with pneumonia but without asthma” (Drabiak K, Kyzer S, Nemov V, El Naqa I, *AI and machine learning ethics, law, diversity, and global impact*, Br J Radiol (2023) 10.1259/bjr.20220934, pág. 2).

²⁵ Os sistemas de IA utilizados na área da medicina são classificados pelo Regulamento da Inteligência Artificial como de *alto risco*, exigindo, por isso, para a sua colocação no mercado, o cumprimento de rigorosas obrigações de transparência, avaliação de conformidade e mitigação de riscos [artigos 6.º, n.º 1 e 9.º do Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024].

²⁶ Cf. Decreto-Lei n.º 29/2024, de 5 de abril, que assegura a execução, na ordem jurídica interna, do Regulamento (UE) 2017/745, relativo aos dispositivos médicos.

proteção” são de particular importância no campo da medicina: o respeito pelos direitos fundamentais implica uma obrigação do Estado de regular a prestação de serviços de saúde de tal forma que sejam tomadas precauções contra danos graves à saúde ocasionados por serviços deficientemente prestados.

Partindo desta base, o Estado deve, por exemplo, obrigar os prestadores de cuidados de saúde a implementar medidas de garantia de qualidade e a respeitar o devido padrão de cuidado. Os direitos fundamentais constituem uma estrutura jurídica vinculativa a que a utilização da IA na medicina se sujeita, a qual não releva apenas para o Estado, mas também para todos os *developers* e prestadores de IA médica.

A todo o ser humano pertence, portanto, o direito de exigir ao Estado que respeite e proteja a sua individualidade também diante de atos de terceiros. Por maioria de razão, o ser humano nunca deve ser completa ou irrevogavelmente submetido a decisões tomadas por sistemas técnicos. Como a IA funciona com base em correlações, as suas mais complexas aplicações devem, em particular, ser sempre monitorizadas por seres humanos para garantir que nenhuma característica especial do pensamento ou da tomada de decisão humana se perda. As decisões de IA médica exigem, portanto, supervisão humana antes que qualquer ação significativa seja tomada com base nela ²⁷. Para já e até ver.

Existe um vasto conjunto de preceitos legais relativos a direitos fundamentais que supõem e exigem a igualdade perante a lei e a não discriminação. A partir deles podem deduzir-se outras condições a estabelecer para o desenvolvimento e funcionamento da IA médica: não só os dados de “entrada” devem ser cuidadosamente verificados quanto à presença de enviesamentos, mas também a operação da IA deve ser constantemente monitorizada. Se a IA médica for aplicada a determinados grupos da população que não foram adequadamente representados nos “dados de

²⁷ “Every human being therefore has a right that the state respects and protects his or her individuality, also towards third persons. Humans must therefore never be completely or irrevocably subjected to technical systems, which also applies to the use of medical AI. Since AI works on the basis of correlations, complex AI applications in particular must always be monitored by human beings to ensure that they do not miss any special features of human thinking or decision-making. From this, it can be deduced that the demands for human oversight expressed in computer science are also required by EU fundamental rights. Decisions of medical AI require human assessment before any significant action is taken on their basis” (David Schneeberger – Karl Stöger – Andreas Holzinger, *The European legal framework for medical AI*, 4th International Cross-Domain Conference for Machine Learning and Knowledge Extraction, Aug 2020, Dublin, Ireland, pág. 2).

treino”, o seu emprego deve também ser questionado de forma particularmente crítica. Ao mesmo tempo, no entanto, cabe garantir que a IA médica útil possa ser disponibilizada a estes grupos da melhor forma possível. Por outras palavras: em homenagem ao princípio da igualdade, a IA médica deve encontrar-se acessível a todos. Mas, para evitar discriminações negativas, a diversidade de pessoas deve ser sempre tida em conta, quer na programação, quer na aplicação.

Por outro lado, os direitos fundamentais não só estabelecem limites à utilização da IA, como podem também promovê-la. Se uma aplicação de IA médica cumprir os requisitos indispensáveis, poderá tornar-se obrigatória a sua utilização. O que inclui o dever de proibir os métodos de tratamento médico que já não possam ser fornecidos com a qualidade exigida sem o envolvimento da IA.

§ 9. Embora a lei seja geralmente flexível quando aborda o impacto de novas tecnologias como a IA, a questão relativa à determinação dos responsáveis por danos originados pela sua utilização não pode ignorar-se ou desleixar-se.

Questões pertinentes são particularmente perturbantes. Quem será legalmente responsável quando a IA médica apresentar imperfeições? O programador de *software*, o fabricante, o pessoal de manutenção, o fornecedor de TI, o hospital, o clínico ²⁸?

Para além da responsabilidade baseada na culpa, existe também, para áreas específicas, a responsabilidade objetiva. A ideia geral a ela subjacente é a de que, embora o uso do objeto perigoso seja socialmente aceitável ou mesmo desejável, o seu operador deve arcar com a obrigação de indemnizar danos por ele causados, independentemente de qualquer falha da sua parte. A responsabilidade baseada na culpa e a responsabilidade objetiva são suscetíveis de sobreposição e funcionamento paralelo.

As regras relativas a ambas – sobretudo as que versam sobre responsabilidade objetiva – destinam-se, em geral, a obter um equilíbrio entre a inovação tecnológica e os riscos potenciais ligados à sua utilização. Em teoria, o sistema de responsabilidade civil proporcionará, assim, incentivo para a criação e manutenção de um padrão de cuidado que priorize

²⁸ “Liability concerns present significant challenges to adoption. As existing and new oversight models develop health care AI systems, the developers of such systems will typically have the most knowledge of risks and be best positioned to mitigate the risk. As a result, developers of health care AI systems and those who mandate use of such systems must be accountable and liable for adverse events resulting from malfunction(s) or inaccuracy in output” (WMA Statement on Augmented Intelligence in Medical Care – Adopted by the 70th WMA General Assembly, Tbilisi, Georgia, October 2019, Challenges).

a segurança do doente, dissuada a introdução de tecnologias excessivamente arriscadas e promova a responsabilização.

Considerando a complexidade e a opacidade dos algoritmos de IA, é questionável se o ónus da prova da causalidade deve recair sobre o lesado. Será difícil rastrear os danos até ao comportamento humano e estabelecer um nexo causal ²⁹. A observância das regras normais sobre a repartição do ónus da prova pode facilmente constituir um obstáculo quase intransponível. Devido a isto, deverá sustentar-se que é ao utilizador de IA – o operador, o produtor, o *designer* ou o prestador de serviços (no caso: o médico ou o hospital) – que cabe suportar o ónus probatório. Esta abordagem é muito provavelmente capaz, contudo, de sufocar a inovação.

§ 10. A falta de consentimento informado (a manifestar pelo paciente) é também suscetível de suscitar responsabilidade. Daí que os deveres de fornecer informações e respeitar a autonomia do paciente sejam parte integrante da medicina contemporânea (e do Direito Médico).

Isto resulta da conceção segundo a qual a tomada da decisão terapêutica há-de ser “partilhada” entre médico e paciente, tendo este último a palavra final. Doente e médico são vistos, por isso, como parceiros iguais. Ao primeiro reconhece-se autonomia para tomar uma decisão livre e informada. O que implicitamente supõe informação suficiente. O dever de a fornecer em relação à utilização e possíveis maus funcionamentos da IA médica dependerá do risco que lhe estiver associado. O que significa que o doente não deve necessariamente ser informado sobre cada utilização específica da IA.

Embora se possa argumentar que não existe um dever geral de divulgar o uso da IA médica junto dos doentes, é possível identificar dois fatores capazes de estabelecer a obrigação inversa: primeiro, a sua já referida opacidade e, segundo, o (aumento) do risco operacional ligado a um sistema de IA.

De uma forma pragmática, é possível sustentar-se que os procedimentos médicos, com ou sem o uso de IA, já atingem hoje uma tal complexidade que a obrigação da equipa médica de os descrever e explicar apropriadamente atingirá reduzida intensidade. Consequentemente, a exigência de consentimento informado não requer que uma decisão de IA

²⁹ “Successful tort claims against a medical AI developer face numerous challenges, including identifying the injury, demonstrating causation within the context of opaque recommendations, overcoming the reluctance of courts to find liability for software generally” (W. Nicholson Price, *Medical AI and Contextual Bias*, Harvard Journal of Law & Technology, Volume 33, Number 1, 2019, pág. 86).

seja compreensível em detalhe ou que se forneça uma extensa elucidação sobre o modo como ela foi especificamente alcançada. De acordo com esta abordagem, bastará que se faculte informação, em termos aproximados, tanto sobre o funcionamento de uma aplicação de IA médica, como sobre o tipo de erros que normalmente se associam à sua operação.

Mas também é possível afirmar, em sentido inverso, que apenas quando a classe médica se encontrar totalmente satisfeita acerca do modo de laboração dos dispositivos de IA ao seu dispor lhes deve dar utilização. De acordo com esta abordagem, apenas se autorizará o recurso à IA médica quando os pacientes tiverem sido esclarecidos sobre as suas funções essenciais, de antemão e de forma inteligível.

Na vida real, a escolha de um prestador de cuidados de saúde e de um método de tratamento dependerá principalmente de aspetos altamente subjetivos (por exemplo, a instituição hospitalar, o fabricante do sistema de IA ou a identidade do médico) e não do conteúdo objetivo da informação fornecida. No entanto, como as impressões pessoais podem ser enganadoras, o municiamento de factos cientificamente sólidos como base para a tomada de decisões informadas continua, de um ponto de vista jurídico, a ser uma pedra angular do conceito de consentimento informado. É certo que ele tem sido algo idealizado e por isso se caracteriza, por vezes, como desnecessariamente complexo e praticamente inatingível. No entanto, esta crítica não se limita à utilização da IA na medicina, mas faz parte da discussão geral sobre os prós e os contras do consentimento informado. Além disso, ele não é um conceito monolítico, apresentando antes diferentes manifestações. Normalmente, as informações necessárias devem ser fornecidas pelo profissional de saúde numa consulta pessoal. No caso dos medicamentos, porém, uma bula é (para além da prescrição) considerada satisfatória para dar cumprimento ao dever de informação. Daí que talvez o fornecimento de uma comparável “bula de IA” se possa visualizar como um meio capaz de dar cumprimento aos padrões de transparência necessários para o recurso à IA em contexto médico.

§ 11. O prestador de serviços médicos (médico, hospital) pode ser responsabilizado pelo tratamento errado ou pela ausência de consentimento informado. Mas, não é responsável pelo sucesso do tratamento. Em geral, a sua obrigação é de meios, não de resultado.

Em tese, a responsabilidade pela utilização de dispositivos médicos dotados de IA poderá despontar, designadamente, quando a recomendação de IA esteja incorreta ³⁰. A opacidade dos algoritmos de IA é sus-

³⁰ “The fundamental challenge of AI from a medical liability standpoint is assessing

cetível, contudo, de inviabilizar a avaliação sobre: (i) a razoabilidade da decisão de utilização da IA; (ii) e a justificação do médico (em função das *legis artis*) para se desviar de uma recomendação de IA ³¹. O que suscita um difícil problema probatório.

Tratando-se de responsabilidade civil é, geralmente, ao lesado que cabe fazer a prova da adequação entre a conduta e o dano. Quando, todavia, se trate de dispositivos dotados de IA, a complexidade e, sobretudo, o *black box effect* torna legítimo questionar se o ónus da referida prova há-de recair sobre ele. É que ser-lhe-á difícil radicar os danos no comportamento humano que conduziu à criação da IA e estabelecer um nexo de causalidade. Pode conjecturar-se, por isso, que há-de ser o operador e/ou o produtor (por exemplo, o *designer*, o prestador de serviços médicos: médico, hospital, etc.) a suportar o ónus da prova (da inadequação entre IA e dano). Ou, indo mais longe, é igualmente concebível a introdução de um regime de responsabilidade contra a própria IA, personificando-a. Pela possível sobrecarga económica (traduzida na multiplicação de indemnizações) associada a uma ou outra solução, elas apresentam a óbvia desvantagem de poderem colocar um freio à inovação e ao progresso tecnológico.

Também é certo, por outro lado, que sempre que juridicamente se torna necessário lidar com problemas novos, é de uso colocar a questão da analogia, da proximidade, com outros que, entretanto, já se encontram em estado mais avançado de resolução.

No que toca a aplicações de IA em objetos que nos rodeiam, crê-se que, por a respetiva tecnologia estar mais desenvolvida e ser mais confiável, é na legislação sobre veículos automóveis que a parecença se há-de estabelecer.

liability when some error and injury inevitably occurs. For example, what happens when an AI algorithm recommends one course of action but a physician believes another course of action, based on their own substantial expertise, is more prudent?8 What happens if the AI recommends a course of treatment and the physician adopts the recommendation, but the AI was wrong and the patient is injured?" (Scott J. Schweikart, *Who Will Be Liable for Medical Malpractice in the Future? How the Use of Artificial Intelligence in Medicine Will Shape Medical Tort Law*, Minnesota Journal of Law, Science & Technology, Volume 22, 2021, pág. 4).

³¹ "Two factors reduce lay judgment of liability: following standard care and following the recommendation of AI tools. These results provide guidance to physicians who seek to reduce liability, as well as a response to recent concerns that the risk of liability in tort law may slow the use of AI in precision medicine" (Kevin Tobia – Aileen Nielsen – Alexander Stremitzer, *When Does Physician Use of AI Increase Liability?*, The Journal of Nuclear Medicine, Vol. 62, No. 1, January 2021, pág. 20).

Pode dizer-se, para inviabilizar a analogia, que a utilização de sistemas de IA médica envolve um número de potenciais intervenientes que, por comparação com o caso dos automóveis autónomos, será muito maior. Neste último teremos: o fabricante de automóveis que desenvolveu e implementou a IA no veículo; depois, o comprador/utilizador; um conjunto de outros compradores/utilizadores; e, por fim, os peões e demais utentes da via pública. É um ecossistema complicado, mas muito menos complexo, diz-se, do que aquele que se encontra na IA médica: os doentes cujos dados são utilizados para construir a IA médica, com a consequente necessidade de consentimento para a sua utilização; as entidades com quem se procede à respetiva partilha; as entidades a quem incumbe estabelecer regras antidiscriminação sobre o respetivo uso; os *developers* a quem pertence tomar decisões centrais sobre o *design* da IA médica, incluindo os conjuntos de dados de treino, os resultados a atingir, os modos de interação, etc.³²; os reguladores que devem determinar em que circunstâncias procederão à revisão da IA médica, a que padrões de segurança ou eficácia ela deve obedecer e se recorrerão a algoritmos “bloqueados” ou “adaptáveis”; e, em sistemas de saúde mais privatizados, os hospitais que decidam sobre que IA médica utilizar.

Não parece, contudo, que a pretensa diferença quantitativa seja particularmente relevante. Justamente ao contrário, até se afigura que, entre as consequências prejudiciais associadas ao uso de automóveis autónomos e aquelas que se ligam ao recurso a dispositivos médicos dotados de IA, o número de potenciais lesados será abstratamente igual: doentes todos estaremos algum dia, utentes da via pública todos somos (seja a que título for).

§ 12. O uso da IA no âmbito da prestação de cuidados de saúde envolve o recurso à chamada análise preditiva (*predictive modeling*).

Dois assuntos principais dominam a controvérsia envolvida na sua utilização: *precisão* e *legitimidade*. Com efeito, pode discutir-se, primeiro, se elas são suficientemente exatas para justificar o seu uso; e, depois, se são legítimas, ao menos em certos contextos, designadamente por se poderem mostrar incompatíveis com o respeito devido à autonomia individual³³.

³² Barry Solaiman and I. Glenn Cohen, *A framework for health, AI and the law*, Research Handbook on Health, AI and the Law, Elgar Publishing, 2024, pág. 4.

³³ “The essential point is that you are not free when you are subject to the arbitrary will of another. Predictive systems are no exception. Predictive systems violate your freedom when they push you down na arbitrary and capricious path, and – importantly for our purposes – they also violate your freedom when you are left with no practical

De um modo geral, a escolha dos preditores em que a análise preditiva há-de assentar orienta-se por critérios de precisão. Mas esta não compõe o único critério subjacente. Em primeiro lugar, por ser possível dar-se o caso de não existir exclusivamente um que satisfaça o mesmo propósito; outras opções haverá igualmente dotadas de suficiente exatidão. Além disso, tendo em consideração que mais objetivos se apresentarão, pode revelar-se importante a escolha de um esquema preditivo que sacrifique algum grau de rigor.

Os modelos de análise preditiva são criados para avaliar dados, descobrir padrões, analisar tendências e aproveitar esse *insight* para antecipar propensões futuras ³⁴. De um modo geral, eles aproveitam os dados históricos para prever eventos vindouros, eventualmente permitindo tomar decisões proativas neles baseadas. Por seu intermédio, descobrem-se padrões e tendências ocultas e fornecem-se informações para o planeamento estratégico e para a eficiência operacional. O que, bem usado, é proveitoso. Por exemplo, de forma muito notória, determinando tratamentos de saúde ³⁵.

§ 13. O que sobremaneira importa sublinhar, contudo, é que as máquinas dotadas de IA se assemelham, em vários aspetos ligados à racionalidade, ao ser humano. Seja qual for a arquitetura técnica que a engenha-

alternative but to submit to the decisions without knowing whether there are adequate reasons for them, reasons that at a minimum show that the decisions are not arbitrary and capricious. You are not free if you are subject to the will of another and denied knowledge of whether that will is arbitrary and capricious. Respecting human freedom requires meeting, or at least sufficiently closely approximating, the following *Knowledge Condition*: those subject to decisions of another are able with reasonable effort to know that there is an adequate justification for the decisions" (Robert H. Sloan – Richard Warner, *Algorithms and Human Freedom*, Santa Clara High Technology Law Journal, vol. 35, 2019, pág. 4).

³⁴ "Predictive analytics is the use of use of mathematical, statistical, and artificial intelligence techniques for classification and prediction" (Robert H. Sloan – Richard Warner, *Algorithms and Human Freedom*, Santa Clara High Technology Law Journal, vol. 35, 2019, pág. 3).

³⁵ "As one noted health law scholar has written, many consider technology pivotal to addressing issues in health care because technology's rapid progress has facilitated affordable and easy access to research advancements and patient records. Given these benefits, the use of Big Data in health care seems particularly attractive. Some even predict that it will bring forth «personalized medicine» – analyzing large data sets of patient information to make diagnostic predictions and treatment suggestions" (Kate Crawford – Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, Boston College Law Review, vol. 55, 93, 2014, pág. 102).

ria informática adote, faz já parte da realidade, do presente, a existência de máquinas dotadas de *deep learning* ³⁶. Desde computadores com intervenção em certos modelos de contratação eletrônica até aos veículos automóveis – exemplo, para já, paradigmático – ou navios de condução completamente autónoma ³⁷, passando por sistemas informáticos capazes de atuar como mediadores em *online dispute resolution* ³⁸ ou como guardas prisionais ³⁹. E adivinha-se, com alto grau de certeza, que o futuro não muito longínquo há de passar pela ampla utilização destes engenhos ⁴⁰.

³⁶ “Deep Learning é um subcampo do Machine Learning (...). A principal característica do deep learning é o uso de redes neurais artificiais profundas, compostas por múltiplas camadas de processamento (daí o termo «profundo»). Inspirado na estrutura do cérebro humano, esse método busca modelar algoritmos que aprendem representações de dados em múltiplos níveis de abstração. Ou seja, ele permite que um sistema aprenda diretamente a partir de dados brutos – como imagens, textos e sons – sem necessidade de pré-processamento ou extração manual de características” (ChatGPT versão base, consultado em 22/04/2025).

³⁷ “Self-driving cars are being released on the roads; we rely on medical diagnosis tools to catch rare diagnoses; product recommendation systems use pattern recognition software to analyze our needs and optimize our shopping experience; we use automated surveillance techniques, killer bots, and other weaponized AI technologies to defend our countries; powerful data mining applications allow us to sift through a wealth of information within the shortest period of time; and AI-enabled decision-making systems using predictive analytics are employed in financial services to detect fraud, tax evasion, or money laundering, and in the justice system for predictive policing and sentencing” [Olivia J. Erdélyi – Judy Goldsmith, *Regulating Artificial Intelligence Proposal for a Global Solution*, 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES ‘18), February 2–3, 2018, New Orleans, LA, USA, pág. 1].

³⁸ “The main reason against sitting computers in the chairs of judges and attorneys, mainly uttered by lawyers, is that doing so is morally undesirable. However, that alone would not hold back the research being done in the area of AI and The Law; it would, at most, delay its implementation but not its development” (Davide Carneiro – Paulo Novais – Francisco Andrade – John Zeleznikow – José Neves, *Online dispute resolution: an artificial intelligence perspective*, Artificial Intelligence Review, 2014, 41, págs. 211–240, pág. 230).

³⁹ “South Korea has the world’s first robot prison guard that can read emotions, identify abnormal prisoner behavior, and alert human authorities to possible concerns. The robot uses 3D cameras and software to study human behavior, make predictions, and then use its wireless communication system to allow human guards to speak to prisoners” (Melanie Reid, *Rethinking the Fourth Amendment in the Age of Supercomputers*, Artificial Intelligence, and Robots, West Virginia Law Review, vol. 119, 2017, pág. 865).

⁴⁰ Inclusivamente no campo das profissões jurídicas (cf. Edwina L. Rissland, *Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*, The Yale Law Journal, vol. 99, 1990, págs. 1961 a 1979). “Attorneys combine their judgment, training, reasoning, analysis, intuition, and cognition under the facts to make approximate legal predictions

Daí a pertinente pergunta: por eventuais danos associados a atividades (condutas?) de máquinas ditas inteligentes, quem responderá (contratual ou extracontratualmente)? A pessoa – singular ou coletiva – que, em benefício próprio, delas se aproveite ⁴¹? A própria máquina ⁴²? O seu produtor ou fabricante?

§ 14. Para conjecturar a hipótese de responsabilização da própria máquina torna-se necessário reconhecer-lhe personalidade jurídica. Ora, qual a razão de ser para colocar esta questão, aparentemente extravagante-

for their clients. To some extent, machine learning algorithms could perform a similar but complementary role, only more formally based upon analyzed data” (Harry Surden, *Machine Learning and Law*, Washington Law Review, vol. 89, 2014, pág. 105). “Artificial intelligence can now engage in legal reasoning, because a well-designed program can tell a lawyer, or even a judge, what cases are really closest to the case at hand, and what cases are properly distinguished from it” (Cass R. Sunstein, *Of Artificial Intelligence and Legal Reasoning*, Public Law & Legal Theory Working Papers n.º 18, University of Chicago, 2001, pág. 5).

⁴¹ “We must assign responsibility for injuries – whether physical, economic, or emotional – caused by interactive and/or self-learning systems. We might hold many different potential actors liable, including the owner, operator, retailer, hardware designer, operating system designer, or programmer(s), to name only a few possibilities.

What degree of fault should we require of these potential defendants? The easy case, and the one for which the law is already well prepared, involves a defendant who programs a robot to harm another. But in most cases, it will be difficult to show either a deliberate intent to harm or knowledge that harm will occur. One might then turn to theories of negligence. But although the risk of some kind of injury at some point in the future is foreseeable whenever one introduces a new technology, how and when an injury occurs may not be particularly foreseeable to each of the potential defendants in the above list. If the law hopes to assign responsibility to humans and corporations, injuries by robotic and AI systems may strain traditional concepts of foreseeability” (Jack M. Balkin, *The Path of Robotics Law*, California Law Review Circuit, vol. 6, 2015, pág. 52).

⁴² “Would it be fair to hold liable the companies that designed, programmed, or manufactured HAL 9000” (o super computador protagonista do filme 2001: *Odisseia no Espaço*), “even though they embedded in HAL’s «thinking» systems the first rule of autonomous machines – i.e., never harm a human – and even though the evidence strongly suggests that HAL «taught» himself to defy their instructions? Or should the creators of machines that have the capacity to «think» be held strictly liable whenever anything goes wrong? If so, on what theory? The theory that the wrongful conduct itself is proof of a defect? Or on an insurance-based theory that the creators are in a better economic position to absorb the cost of the injury than the person harmed?” (David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, Washington Law Review, vol. 89, 2014, pág. 125).

te ⁴³? “Can there be such a thing as AI «free will» ⁴⁴? “Will the agents designed be autonomous enough to deserve and warrant legal rights” ⁴⁵? Poderá sustentar-se que “*computo, ergo sum*”?

A partir do instante em que uma máquina se encontre dotada de *deep learning*, pode já falar-se, com alguma propriedade, de inteligência e de capacidade para tomar deliberações autônomas ⁴⁶, para decidir (ainda que dentro de estreitas fronteiras). E tais qualidades são suscetíveis de produzir a equiparação (com maior ou menor proximidade) entre máquina e ser humano ⁴⁷. Pelo que, portanto, pode, na atuação jurídica, discutir-se a imputação da autoria de condutas.

⁴³ Não tanto assim, no entanto, quando se leve em consideração que a “English admiralty law treats a ship as a legal person capable of being sued in its own right; other legal systems have recognized temples, dead persons, spirits and idols as legal persons” (Samir Chopra – Laurence White, *Artificial Agents – Personhood in Law and Philosophy*, ECAI’04 Proceedings of the 16th European Conference on Artificial Intelligence, pages 635-639, pág. 635).

⁴⁴ Sam N. Lehman-Wilzig, *Frankenstein Unbound – Towards a legal definition of Artificial Intelligence*, Futures, vol. 13, Issue 6, December 1981, pág. 452.

⁴⁵ Samir Chopra – Laurence White, *Artificial Agents – Personhood in Law and Philosophy*, ECAI’04 Proceedings of the 16th European Conference on Artificial Intelligence, pages 635-639, pág. 635.

⁴⁶ “What attributes of particular agents should the legal system take into consideration in determining whether to grant personhood to particular agents? The practical capacity to perform cognitive tasks will presumably be of primary importance” (Samir Chopra – Laurence White, *Artificial Agents – Personhood in Law and Philosophy*, ECAI’04 Proceedings of the 16th European Conference on Artificial Intelligence, pages 635-639, pág. 639).

⁴⁷ “There are five attributes that one would expect an intelligent entity to have. The first is communication. One can communicate with an intelligent entity. The easier it is to communicate with an entity, the more intelligent the entity seems. (...) The second is internal knowledge. An intelligent entity is expected to have some knowledge about itself. The third is external knowledge. An intelligent entity is expected to know about the outside world, to learn about it, and utilize that information. The fourth is goal-driven behavior. An intelligent entity is expected to take action in order to achieve its goals. The fifth is creativity. An intelligent entity is expected to have some degree of creativity. In this context, creativity means the ability to take alternate action when the initial action fails. A fly tries to exit a room and bumps into a windowpane continues to repeat the same futile behavior. When an AI robot bumps into a window, it tries to exit using the door. Most AI entities possess these five attributes by definition. Some twenty-first century types of AI entities possess even more attributes that enable them to act in far more sophisticated ways” (Gabriel Hallevy, *The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control*, Akron Intellectual Property Journal, vol. 4, 2010, págs. 175/176).

Se v.g. o veículo de condução autónoma for confrontado, no seu movimento, com uma situação em que, para evitar embater num ciclista que se desequilibra à sua frente, o desvio de rota implica atropelar uma criança, qualquer que seja a opção que tome, defrontará, antes de mais, a necessidade de proceder à escolha ⁴⁸. Independentemente do critério a que recorra, mesmo que ele consista na pura casualidade (que ainda é um critério), o dilema pôr-se-á e deverá ser instantaneamente resolvido pela própria máquina ⁴⁹. De acordo com a primeira lei de Asimov ⁵⁰, uma máquina deve ser programada para preservar sempre a vida humana. Mas como proceder quando guardar uma vida humana envolva o sacrifício de outra? A escolha há de forçosamente ser deixada ao critério da máquina, ao menos quando não haja oportunidade de, em tempo útil, recorrer à decisão de um ser humano ⁵¹.

O perigo posto pela IA situa-se, não na génese da máquina e nas tarefas que esta potencialmente possa desempenhar, mas sim na aptidão

⁴⁸ Cf. Amitai Etzioni – Oren Etzioni, *Incorporating Ethics into Artificial Intelligence*, Journal of Ethics, vol. 17, n.º 3, 2013.

⁴⁹ “Driverless cars, viewed as the archetypal autonomous machines, are learning machines. They are programmed to collect information, process it, draw conclusions, and change the ways they conduct themselves accordingly, without human intervention or guidance” (Amitai Etzioni – Oren Etzioni, *Incorporating Ethics into Artificial Intelligence*, Journal of Ethics, vol. 17, n.º 3, 2013). “Like human drivers, the machines that drive Google cars will on occasion encounter unexpected events that call for snap judgments: a child darting in front of a car; a tree limb crashing down just a few yards ahead; a car running a red light; or a patch of black ice that is undetectable on a moonless night” (David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, Washington Law Review, vol. 89, 2014, págs. 126/127).

⁵⁰ As chamadas *Three Laws of Robotics* (Isaac Asimov, *Runaround*, I, Robot, Gnome Press, 1950, pág. 27) são: “*First Law* – A robot may not injure a human being or, through inaction, allow a human being to come to harm. *Second Law* – A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. *Third Law* – A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws”.

⁵¹ As diretrizes contidas no artigo 8 da proposta de Carta de Direitos Fundamentais Digitais da União Europeia (ZEIT – Stiftung Ebelin und Gerd Bucerius) – “(1) Decisions which have ethical implications or which set a precedent may only be taken by a person” e “(3) A natural or legal person must always be responsible for the actions of self-learning machines and their consequences” –, a serem aprovadas e colocadas em vigor tal qual, valerão o que puderem valer. A única maneira de lhes dar plena execução consiste em proibir a IA. Caso contrário, forçar-se-ão muitas soluções fantasiosas, desligadas da realidade.

que os *softwares* dotados de IA têm para “em algumas situações, modificar as instruções que lhes foram dadas, levando a cabo atos que não estão de acordo com uma programação pré-definida, mas que são potenciados pela integração com o meio” ⁵².

A *autonomia* que a caracteriza – fundada na sua capacidade para a autoaprendizagem – conduz à *imprevisibilidade* (ou previsibilidade limitada) da respetiva atuação ⁵³.

De um modo geral, a personificação de realidades não humanas tem sido principalmente entendida “as a strategy of dealing with the uncertainty about the identity of the other” ⁵⁴. Mas, no caso das máquinas munidas de *deep learning*, a justificação para a personificação é mais forte. Ao encontrar-se dotada de capacidade de deliberação, ela formará mais provavelmente um *moral agent* do que *v.g.* um animal. É por isso concebível adstringi-la a deveres ⁵⁵. O que viabilizará, por exemplo, a sua vinculação, no plano da responsabilidade civil (extracontratual ou contratual), à obrigação de indemnizar ⁵⁶.

⁵² Mafalda Miranda Barbosa, *Inteligência artificial, E-Persons e Direito: Desafios e Perspetivas*, in Direito e Robótica, Instituto Jurídico Centro de Direito do Consumo – Faculdade de Direito da Universidade de Coimbra, 2020, pág. 59.

⁵³ “In the case of advanced AI, a system that learns from environmental data may act in ways that its designers have no feasible way to foresee” (Peter M. Asaro, *The Liability Problem for Autonomous Artificial Agents*, in 2016 AAAI Spring Symposia, Stanford University, AAAI Press, 2016, pág. 191).

⁵⁴ Gunther Teubner, *Rights of Non-humans? – Electronic Agents and Animals as New Actors in Politics and Law*, Journal of Law and Society, vol. 33, n.º 4, 2006, pág. 497.

⁵⁵ “If AI would be fully autonomous (such as superintelligence), then they must be aware of their actions. If they are aware of their actions, they must be liable for their actions. AI’s autonomy in law means that AI has rights and a corresponding set of duties. In law, rights and duties are attributed to legal persons, both natural (such as humans) and artificial (such as corporations). Therefore, if we seek for AI to be liable for its actions, we should ascribe legal personhood to it” (Paulius Cerka – Jurgita Grigien – Gintar Sirbikyt, *Liability for damages caused by artificial intelligence*, Computer Law & Security Review, n.º 31, 2015, pág. 383).

⁵⁶ “Usually... people make a whole range of anthropomorphic assumptions about non-humans and they act accordingly, as if they were humans. The non-humans are supposed to process meaning self-referentially as the humans do, to be equipped with freedom of choice, with self-preserving strategies, reflective capacities, phenomenal world views of their own, empathy and understanding, even with the ability to communicate. (...)

In this view, there is no difference between human and non-human actors, strange as this may sound. «Personality means nothing but the symbolic signification of the capacity

§ 15. A responsabilização por danos causados a outrem pode dar-se *objetivamente* – pelo risco ou pelo sacrifício (artigo 483.º, n.º 2, Cód. Civil). Será por esta via possível vincular uma máquina à obrigação de indemnizar, independentemente de ela compor um *moral agent* e, portanto, ainda que não disponha de capacidade para se nortear no sentido requerido pela imposição de um certo dever jurídico? De acordo com a máxima *cuius est commodum eius est periculum*, importa, sobretudo, que isso se apresente como uma solução justa. O juízo de culpa e os respetivos pressupostos tornar-se-ão aqui prescindíveis.

Frente a esta solução, um obstáculo de distinta natureza pode, porém, vislumbrar-se.

Embora nem sempre se lhe conceda suficiente ênfase, o reconhecimento de personalidade jurídica envolve a presença de um património. O incumprimento de deveres acarreta, como regra, a responsabilização do sujeito que nele haja incorrido. E, frequentemente, se daí emergirem danos para outra pessoa, sobrevirá a obrigação de os reparar (indemnizando ou compensando). Razão para se prescrever que “pelo cumprimento da obrigação respondem todos os bens do devedor suscetíveis de penhora” (artigo 601.º, Cód. Civil).

Ora, como poderá uma máquina encontrar-se munida de um património passível de, em caso de incumprimento de algum dever, lhe permitir responsabilizar-se por danos ocasionados à pessoa ou ao património de outrem?

§ 16. Quando se entenda que a máquina é pessoa jurídica e que, por isso, se torna suscetível de titular direitos e deveres – com isso se viabilizando a possibilidade de a obrigar a reparar os desmandos feitos a direitos alheios –, pode ainda questionar-se se a pessoa que dela se serve e/ou extrai benefícios – utente, empresário, produtor, etc. – não deverá ficar, com fundamento na ideia de *respondeat superior*⁵⁷, solidariamente

to participate in communication, and it does not matter, and it is historically variable whether the relevant entities are gods, animals, spirits, robots or humans” (Gunther Teubner, *Rights of Non-humans? – Electronic Agents and Animals as New Actors in Politics and Law*, *Journal of Law and Society*, vol. 33, n.º 4, 2006, pág. 504).

⁵⁷ “On the issue of *respondeat superior* (*Erfüllungsgelhilfe*, *Verrichtungsgelhilfe*) three positions are debated: apart from the responsibility for human errors, computer mistakes are seen by some authors as *force majeure* without any liability of the owner. Others suggest strict liability for computer mistakes or a contractual guarantee as an implied condition of contract. They treat artificial agents as mere tools of their operators, or as mere means of communication. All actions of artificial agents are attributed to the agent’s operator. They come up with a stricter liability principle than that which applies to human agents and

vinculada. Designadamente, a título de comitente ⁵⁸.

Se o comissário for máquina, que vínculo se há de conjecturar existir com o comitente?

Na visão normal, a máquina é um escravo ⁵⁹. É uma coisa, como tal desprovida de dignidade, não merecendo, por isso, qualquer especial respeito. Mesmo aceitando esta conceção (que se adivinha não poder permanecer por muito tempo) ⁶⁰, há uma objeção que não se pode descurar. Se a máquina substitui o ser humano, desempenhando funções que antes este cumpria, e se o trabalhador humano tinha, nessa qualidade, direito a remuneração pelo serviço prestado, por que razão a máquina que no seu lugar se sub-roga não há-de a ter? Por que razão caberá desonerar a entidade patronal (ou equivalente) do seu pagamento quando o trabalho que em seu benefício se presta seja executado por uma máquina? Se essa for a exclusiva finalidade de tal sub-rogação, haverá ela de ter-se como meritória?

Poderá, de imediato, contestar-se que a retribuição salarial é devida com fundamento na necessidade de garantir a subsistência do trabalha-

their principals. Most interesting for our context is a third position which makes an analogy to *respondeat superior* and defines normalised expectations for the technical capacities of computer action. This indeed would be the equivalent of duty of care applicable to human actors and corporate actors" (Gunther Teubner, *Rights of Non-humans? – Electronic Agents and Animals as New Actors in Politics and Law*, Journal of Law and Society, vol. 33, n.º 4, 2006, pág. 509). Cf., a este propósito, o disposto no artigo 33.º do Decreto-Lei n.º 7/2004, de 7 de janeiro: "1 – À contratação celebrada exclusivamente por meio de computadores, sem intervenção humana, é aplicável o regime comum, salvo quando este pressupuser uma atuação. 2 – São aplicáveis as disposições sobre erro: a) Na formação da vontade, se houver erro de programação; b) Na declaração, se houver defeito de funcionamento da máquina; c) Na transmissão, se a mensagem chegar deformada ao seu destino. 3 – A outra parte não pode opor-se à impugnação por erro sempre que lhe fosse exigível que dele se apercebesse, nomeadamente pelo uso de dispositivos de deteção de erros de introdução".

⁵⁸ Cf. António Pinto Monteiro, "*Qui facit per alium, facit per se*" – será ainda assim na era da robótica?, Revista de Legislação e Jurisprudência, Ano 148.º, n.º 4015, Março - Abril 2019, págs. 200-211.

⁵⁹ "The term «robot» stems from the Czech word *robota*, meaning drudgery, servitude, or forced labour. From the beginning, then, its purpose was to function as humanity's modern slave" (Sam N. Lehman-Wilzig, *Frankenstein Unbound – Towards a legal definition of Artificial Intelligence*, Futures, vol. 13, December 1981, pág. 449).

⁶⁰ Levando em conta, por um lado, a procriação artificial, a clonagem, a manipulação genética e hipóteses afins, e, por outro, a outorga, a máquinas munidas de inteligência artificial, da capacidade para formular juízos de valor e para ter sensações ou sentimentos, antecipa-se que, mais cedo do que tarde, a própria separação entre pessoa e máquina seja rápida e profundamente posta em crise.

dor. Mas, sobretudo, que ela se destina – quer efetivamente a assegurar, quer não – a preservar a sua dignidade impedindo a coisificação da pessoa. Só que então que justificação se há de avançar para legitimar o pagamento a uma sociedade de alguma retribuição por serviço prestado (ainda que não se lhe conceda a designação de salário)? As pessoas coletivas consideram-se titulares desta componente da dignidade?

Personificando a máquina, nada parece obstar, pois, a que ela, enquanto comissário, se encontre juridicamente ligada ao comitente mediante contrato de trabalho ou situação jurídica afim.

§ 17. A opacidade (a que antes se aludiu) que muitos sistemas de IA apresentam constitui um problema suplementar ⁶¹. Especialmente, quando o respetivo *firmware* tenha sido desenvolvido ou modificado através de técnicas de *machine learning* ou de *deep learning* ⁶². O que é capaz de suscitar sérios obstáculos probatórios, ainda que de pura ordem prática, no que respeita à causalidade. “Even without changes to the original software design, the embedded criteria steering the collection and analysis of data and the decision-making process may not be readily explicable and often require costly analysis by experts. This may in itself be a primary practical obstacle to pursuing a claim for compensation” ⁶³. E mesmo quando se prove que certo evento desencadeou o dano (por exemplo, o dispositivo médico de

⁶¹ “Modern deep neural networks can be very deep and are extremely interconnected. This means that there may not be any clear way of understanding the decision-making process of the network once it is trained on the data. Moreover, the inputs to machine-learning algorithms, including deep neural networks, are often multi-dimensional, meaning that various input parameters are encoded as high dimensional vectors. Machine-learning algorithms, such as Support Vector Machines, rely on special relationships in higher-dimensional vector spaces. In other words, if there are 115 different parameters used by a model, then machine-learning algorithms will search for patterns in 115 or more dimensions, a sort of geometric space that humans simply cannot visualize. The net effect is both opacity from the vast number of interconnected layers and the difficulty of visualizing higher-dimensional patterns. So, there is no clear way for human beings to easily examine the patterns that a machine-learning algorithm may be seizing on as part of its decision-making process” (Yavar Bathaee, *Artificial Intelligence Opinion Liability*, págs. 141 e 142).

⁶² Daqui em diante acompanha-se o que se disse em José Martinho González, *A responsabilidade civil emergente da detenção de veículos autónomos dotados de inteligência artificial (IA)*, Quid Juris, Lisboa, 2025, págs. 74 a 84.

⁶³ *Liability for Artificial Intelligence and other emerging digital technologies*, European Commission, Expert Group on Liability and New Technologies New Technologies Formation, 2019, pág. 20.

IA que não deteta uma infeção óbvia), a relação causa/efeito nem sempre será igualmente evidente. O referido dispositivo pode ter sido mal projetado (seja no *hardware*, no *software* pré-instalado ou em ambos); pode ter procedido a uma deficiente leitura de dados corretos ou ter recebido dados incorretos; pode uma atualização do *software* feita pelo produtor original ou por terceiros apresentar falhas; ou pode, por fim, dar-se o caso de o usuário não ter instalado uma atualização que teria evitado o erro ⁶⁴.

A conjugação entre opacidade e causalidade alternativa é passível de, por aplicação das regras que tipicamente regem a responsabilidade civil aquiliana, impedir o lesado de obter a devida indemnização. Quando vários sejam os responsáveis e não se consiga determinar, com exatidão, o grau de intervenção de cada qual, uma solução consistirá em vincular todos à obrigação de reparação a título de *concurrent tortfeasors* ⁶⁵. Embora correndo o risco de responsabilizar quem realmente não contribuiu para o dano. Outra maneira, talvez mais eficaz, de obviar o problema consistirá em estabelecer presunções legais de culpa e/ou de causalidade contra o produtor, fabricante, inventor, *developer* ⁶⁶.

§ 18. O instituto da responsabilidade do produtor baseia-se no princípio de que este (definido – ao longo da cadeia de distribuição – de forma ampla: artigo 2.º, Decreto-Lei n.º 383/89, de 06 de novembro) é responsável pelos danos causados ao consumidor pelo defeito de um produto que haja colocado em circulação. Estendida aos sistemas de IA, pressupõe, evidentemente, a ideia de que “AI entities, or at least the majority of them,

⁶⁴ “Liability for Artificial Intelligence and other emerging digital technologies”, European Commission, Expert Group on Liability and New Technologies New Technologies Formation, 2019, p. 22.

⁶⁵ “Sucedee sempre que vários intervenientes coloquem, por si só (ou seja, individualmente), condições suficientes para a produção do – único e mesmo – dano, todos para ele contribuindo (total ou parcialmente), mas cada qual independentemente dos demais” (José A. R. L. González, *Direito da Responsabilidade Civil*, Quid Juris, Lisboa, 2ª edição, págs. 132 e 133).

⁶⁶ Se a questão se puser no âmbito contratual, será sempre possível instituir, por acordo, ao abrigo da liberdade contratual, uma cláusula de inversão do ónus da prova da culpa (para o respetivo conceito, cf. Ana Prata, *Cláusulas de exclusão e limitação da responsabilidade contratual*, Almedina, Coimbra, 1985, págs. 96 e 97) entre o consumidor e o fabricante através da qual o primeiro assumia o encargo de demonstrar positivamente (e não só por presunção – artigo 799.º, n.º 1, Código Civil) a culpa do segundo. Mas, em sede de responsabilidade por danos decorrentes do emprego de sistemas de IA, não é uma solução natural, nem, provavelmente, viável, dada a já muitas vezes aludida *opacidade* que os caracteriza.

are nothing more than toasters or fridges”⁶⁷. São coisas, portanto.

Ora, sucede que, primeiro, o conceito de produto não se afigura claramente suscetível de abarcar organismos de IA⁶⁸; nestes, com efeito, sistemas, produtos e serviços interagem ininterruptamente, tornando-se inexequível uma separação nítida entre si. O que põe em causa o conceito de produto enquanto ele se definir como qualquer *coisa* móvel, ainda que incorporada noutra (artigo 3.º, n.º 1, Decreto-Lei n.º 383/89, de 06 de novembro). Acresce, além disso, que não se afigura nítido que o *software* se deva entender como coisa⁶⁹. É certo que a Diretiva (UE) 2024/2853 do Parlamento Europeu e do Conselho de 23 de outubro de 2024, através do n.º 1 do seu artigo 4.º, vem acrescentar o *software* e os ficheiros de fabrico digitais (mas não a generalidade dos ficheiros digitais) ao conceito de produto. Acolheu-se a ideia segundo a qual as coisas ditas incorpóreas ainda são coisas. O que, todavia, além de não ser incontestável, colide com o preceituado nos artigos 1302.º e 1303.º do Cód.Civil.

Por outro lado, segundo, como a presença de um defeito se mede atendendo, genericamente, às expectativas de segurança do consumidor médio (artigo 4.º, n.º 1, Decreto-Lei n.º 383/89, de 06 de novembro)⁷⁰, a aludida interconexão entre produtos e sistemas torna difícil a sua identificação. Com efeito, a IA mais sofisticada, munida de *deep learning* ou de *machine learning*, faz despontar a questão de saber se desvios imprevisíveis no caminho da tomada de decisão devem ser tratados como defeitos. Mesmo que, com muita benevolência, se possam ver como tal, a complexidade e a opacidade das tecnologias digitais acabam por diminuir consideravelmente, para o lesado, as oportunidades de demonstração do defeito e da causalidade.

⁶⁷ Anat Lior, *AI Entities as AI Agents: Artificial Intelligence Liability and the AI Respondeat Superior Analogy*, pág. 1056.

⁶⁸ Pela mesma exata razão foi necessário editar legislação especial relativa à responsabilidade civil emergente de danos causados por medicamentos. Cf. Vera Lúcia Raposo, *Danos Causados por Medicamentos: Enquadramento Jurídico à luz do Ordenamento Europeu*, Almedina, Coimbra, 2018.

⁶⁹ “For a product liability cause of action, the algorithm, or in some cases the software the algorithm is embedded in, must be a «product» (and not a service). Accordingly, two criteria are often presented. The first is that a product – as distinguished from a service – must consist of some physical embodiment that is available to the purchaser directly. (...) And second, the algorithm must not be a unique or specially designed item. For algorithms designed to control robots which are created in research labs, this second point may be a difficult bar to overcome” (Woodrow Barfield, *Liability for autonomous and artificially intelligent robots*, in *Journal of Behavioral Robotics*, vol. 9, 2018, pág. 197).

⁷⁰ Cf. Calvão da Silva, *Responsabilidade Civil do Produtor*, Almedina, Coimbra, 1999, págs. 633 a 653.

Para determinar a responsabilidade do produtor, terceiro, atende-se ao momento em que o produto foi colocado em circulação (artigo 4.º, n.º 1, Decreto-Lei n.º 383/89, de 6 de novembro). Após este instante, a lei não estabelece, contra ele, quaisquer deveres de monitorização. Daí a razão para que se lhe permita a invocação da chamada *state-of-the-art defence* [artigo 5.º, alínea e), Decreto-Lei n.º 383/89, de 06 de novembro]. Ora, sistemas de IA altamente sofisticados nem podem ser tidos como produtos acabados, nem são colocados no mercado pela forma tradicional. Com efeito, não só o produtor (operador *back-end*) poderá sempre reter algum grau (maior ou menor) de controlo sobre o seu desenvolvimento futuro mediante adições ou atualizações, como pode ele ficar limitado e não ser exclusivo sempre que operar o produto dependa de dados fornecidos por terceiros, de processos de autoaprendizagem ou de configurações personalizadas escolhidas pelo utilizador.

Insistindo-se, para efeitos de indemnização por danos causados, no recurso ao instituto da responsabilidade do produtor, acabar-se-á por sustentar, no fundo, a regra *res ipsa loquitur*: presunção (ainda que refutável) de negligência imputável àquele, assente (i) na prova de que o dano normalmente não teria ocorrido sem ela, (ii) que o objeto que o provocou estava sob o seu controle e (iii) que inexistem outras explicações plausíveis⁷¹. Considerando tudo o que se disse sobre as principais características dos sistemas de IA, trata-se de solução, pelo menos, extraordinariamente forçada.

§ 19. Isto não impediu, no entanto, que a já referida Diretiva (UE) 2024/2853 do Parlamento Europeu e do Conselho de 23 de outubro de 2024, mediante a redação conferida, em especial, à alínea c) do n.º 2 do seu artigo 7.º, a tenha adotado: “Na avaliação da qualidade defeituosa de um produto, devem ser tidas em conta todas as circunstâncias, incluindo: (...) O efeito no produto de qualquer capacidade de continuar a aprender ou de adquirir novas características depois de ser colocado no mercado ou em serviço”.

Esta solução é, desde já, passível de diversas observações críticas.

Primeiro. Dado o altíssimo nível de risco envolvido, uma solução deste tipo irá traduzir-se em potenciais custos proibitivos para o produtor

⁷¹ Para a questão “Who pays?”, na perspetiva da responsabilização do produtor, “the only feasible approach, it would seem, would be to infer a defect of some kind on the theory that the accident itself is proof of defect, even if there is compelling evidence that cuts against a defect theory. There is precedent for courts making such an inference, which is simply a restatement of *res ipsa loquitur*” (David C. Vladec, *Machines Without Principals: Liability Rules and Artificial Intelligence*, in *Washington Law Review*, n.º 117, 2014, pág. 128).

que decida introduzir no mercado qualquer produto dotado de sistemas de IA. Haverá, pelo menos, um (forte) incentivo negativo.

Segundo. Por outro lado:

(i) Se é certo, de acordo com o n.º 1 do artigo 5.º da citada Diretiva, que se considera lesada qualquer pessoa que sofra “danos causados por um produto defeituoso”, logo se acrescenta desde que seja “pessoa singular”. Excluem-se as pessoas coletivas, portanto, do direito à indemnização.

(ii) Mesmo no que toca à pessoa singular, nem todos os danos não patrimoniais são compensáveis. Apenas o dano da morte e a lesão à saúde geral da vítima [artigo 6.º, n.º 1, alínea a)].

(iii) E, ainda quanto à pessoa singular, também nem todos os danos patrimoniais são indemnizáveis [cf. subalíneas i), ii), iii) da alínea b) do n.º 1 do artigo 6.º].

A responsabilidade do produtor está concebida para a indemnização por danos causados ao *utente* do produto defeituoso. Nessa relação, ele assume, tipicamente, a qualidade de *consumidor*^{72 73}. Daí que: (a) as pessoas coletivas fiquem excluídas do seu âmbito de aplicação, dado que elas não têm sido concebidas, na visão mais estrita, como estando dotadas de tal qualidade [artigo 2.º, alínea g), Decreto-Lei n.º 84/2021, de 18 de outubro]⁷⁴; (b) ela também se excetue em relação a “bens utilizados exclusivamente para fins profissionais”.

Ora, o que sobretudo se considera no presente estudo, mas também o que, usualmente, mais importa no que concerne a sistemas de IA, é a res-

⁷² É o que se extrai, por exemplo, do facto de a identificação do defeito se fazer, antes do mais, pela “apresentação e as características do produto, incluindo a sua rotulagem, conceção, características técnicas, composição e embalagem, e as instruções de montagem, instalação, utilização e manutenção” [artigo 7.º, n.º 2, alínea a), Diretiva (UE) 2024/2853 do Parlamento Europeu e do Conselho de 23 de outubro de 2024.

⁷³ Mark A. Geistfeld, *Situating Bystanders Within Strict Products Liability*, in Brooklyn Journal of Corporate, Financial & Commercial Law, vol. 18, Issue 1, The Law of Torts: Duty, Design, & Conflicts, A Festschrift in Honor of Aaron Twerski.

⁷⁴ V.g. acórdão da Relação de Guimarães de 15/12/2022, Proc. n.º 777/20.7T8VCT. G1: “1. A responsabilidade civil do produtor por produtos defeituosos, surge no nosso ordenamento jurídico através do Decreto-Lei n.º 383/89, de 06 de novembro em virtude da transposição da Diretiva n.º 85/374/CEE do Conselho, de 25 de julho de 1985. 2. Este diploma teve como propósito proteger os lesados no âmbito da compra e venda de produtos defeituosos, consagrando o princípio da responsabilidade objetiva do produtor, independentemente de culpa, complementando o direito comum e assegurando desse modo uma maior eficácia na proteção do consumidor (...). 3. O conceito de consumidor reporta-se necessariamente a uma pessoa singular, neste não se incluindo em caso algum as pessoas coletivas, sendo esse o âmbito subjetivo de aplicação do regime previsto no D.L. n.º 383/89 quanto ao lesado”.

ponsabilidade por danos causados a terceiro decorrentes do (ou associados ao) seu uso. No caso de dispositivos médicos dotados de IA, terceiro será quem for paciente e será também a pessoa a quem, neste contexto, mais interessa dar proteção.

É certo que já no domínio da redação inicial do Decreto-Lei n.º 383/89, de 6 de novembro, se defendia que “no dano da morte ou na lesão pessoal, (...) a proteção da saúde e da segurança é a mesma para todos lesados por produto defeituoso circulante no mercado, seja um consumidor propriamente dito, um profissional que utiliza o produto no exercício da sua profissão ou mesmo um simples *bystander*”⁷⁵.

Será útil, por isso, esclarecer o que, para este efeito, cabe entender por terceiro.

Ante o produtor, o consumidor é, quase sempre, um terceiro. Raramente a cadeia de distribuição será tão curta que os interligue, diretamente, por via contratual⁷⁶.

Quando, todavia, se afirma que o produtor se pode encontrar obrigado a indemnizar um *bystander* por danos causados pelo produto defeituoso, joga-se com um distinto conceito de terceiro: *v.g.* outros utentes do produto além do consumidor (familiares, amigos), os passageiros do automóvel, os bombeiros ou a equipa médica que intenta o salvamento do consumidor ou do seu familiar, pessoas vizinhas, etc.

O *bystander*, na lógica do Decreto-Lei n.º 383/89, de 6 de novembro, será, porém, um titular accidental do direito à indemnização: ela é-lhe concedida no pressuposto de os danos eventualmente se *propagarem* contra si como consequência da utilização que o consumidor haja dado ao produto defeituoso. Ora, em especial quando o sistema de IA se torne absolutamente autónomo, a identificação de um utente há-de, quase garantidamente, revelar-se difícil ou impossível tarefa.

Tratando-se de dispositivos médicos dotados de IA, será razoável considerar que o próprio paciente é o consumidor?

§ 20. Em alternativa à responsabilização dos produtores de dispositivos médicos dotados de IA por danos decorrentes do seu emprego, pode conjecturar-se em responsabilizar, antes, os seus usuários.

A cada médico ou profissional equiparado cabe supervisioná-los, prestando atenção quando os utiliza e devendo intervir quando neces-

⁷⁵ Cf. Calvão da Silva, *Responsabilidade Civil do Produtor*, Almedina, Coimbra, 1999, pág. 698.

⁷⁶ Daí a necessidade de conceber a chamada *ação direta* nos termos do n.º 1 do artigo 6.º do Decreto-Lei n.º 67/2003, de 8 de abril.

sário para prevenir acidentes. A responsabilidade do utilizador, quando eles se concretizem, basear-se-á então em descuido, falta de diligência, imperícia, desatenção. Sendo até concebível que legalmente se presuma a respetiva culpa.

Os dispositivos médicos autónomos perderiam assim, no entanto, grande parte da sua utilidade. Não seria viável, por exemplo, pô-los sozinhos em funcionamento para realização de diagnósticos. Mais sério ainda: sendo autónomo ao ponto de a intervenção humana dele se encontrar absolutamente excluída (por nem sequer existirem comandos para o efeito e por ela, a ser exequível, aumentar o perigo de erro em vez de o diminuir), que sentido fará presumir a culpa do utilizador?

Outra opção consistirá em considerar o utilizador do referido dispositivo como responsável pelos danos resultantes de eventuais incidentes, ainda que sobre si não incida o dever de interferir na sua utilização (até por, possivelmente, nenhuma maneira existir de o fazer). Com a seguinte justificação: se assume o risco de usar o objeto, sabendo que ele pode causar danos, aceita os perigos inerentes e a sua provável concretização.