

Universidades Lusíada

Rocha, Leonardo Vieira da

Responsabilidade civil por danos causados por sistemas de inteligência artificial

<http://hdl.handle.net/11067/7869>

Metadata

Issue Date 2024

Abstract Esta dissertação aborda a responsabilidade civil por danos causados por sistemas de Inteligência Artificial (IA), uma questão emergente na era tecnológica. O desenvolvimento tecnológico traz inúmeros benefícios, mas também desafios, especialmente legislativos na área da responsabilidade. Para abordar o tema central, começamos pela determinação do conceito de IA, que, devido à sua complexidade, não tem uma definição unânime, resultando assim em diversas versões conceituais. Analisaremos também o ...

This dissertation addresses civil liability for damages caused by artificial intelligence (AI) systems, an emerging issue in the technological era. Technological development brings numerous benefits but also challenges, especially legislative ones in the field of liability. To address the central theme, we begin by defining the concept of AI, which, due to its complexity, lacks a unanimous definition, resulting in various conceptual versions. We will also analyze the historical framework and som...

Keywords Direito, Responsabilidade civil, Inteligência artificial - Direito e legislação, Inteligência artificial - Aspectos sociais

Type masterThesis

Peer Reviewed No

Collections [ULP-FD] Dissertações

This page was automatically generated in 2025-04-02T23:16:21Z with information provided by the Repository



UNIVERSIDADE LUSÍADA PORTO

**Responsabilidade Civil por danos Causados por Sistemas de
Inteligência Artificial**

Leonardo Vieira da Rocha

Dissertação para obtenção do Grau de Mestre

Porto, 2024



UNIVERSIDADE LUSÍADA PORTO

Responsabilidade Civil por danos Causados por Sistemas de Inteligência Artificial

Leonardo Vieira da Rocha

Dissertação para obtenção do Grau de Mestre

Orientadora: Professora Doutora Cristina Aragão Seia

Porto, 2024

AVISO LEGAL

O conteúdo desta tese reflete as perspectivas, o trabalho e as interpretações do autor no momento da sua entrega. Esta tese pode conter incorreções, tanto conceituais como metodológicas, que podem ter sido identificadas em momento posterior ao da sua entrega. Por conseguinte, qualquer utilização dos seus conteúdos deve ser exercida com cautela. Ao entregar esta tese, o/a autor(a) declara que a mesma é resultante do seu próprio trabalho, contém contributos originais e são reconhecidas todas as fontes utilizadas, encontrando-se tais fontes devidamente citadas no corpo do texto e identificadas na secção referências. O autor, declara, ainda, que não divulga na presente tese quaisquer conteúdos cuja reprodução esteja vedada por direitos de autor ou de propriedade industrial.

Declaração sob compromisso de honra

(Artigo 6.º, n.º 2 das Normas e orientações para a submissão de trabalhos académicos na plataforma Urkund para deteção de similaridade e plágio)

Eu, abaixo assinado, tenho consciência de que a prática de plágio consiste numa forma de violação da integridade académica, constituindo um crime punível por lei com relevância nos regimes disciplinar, civil e criminal.

Nesse sentido, declaro por minha honra que a dissertação/tese apresentada é original e que todas as fontes, incluindo as da minha autoria, estão devidamente identificadas e referenciadas.

Porto, 28 de outubro de 2024

O (A) Estudante

Leonardo Rocha

Agradecimentos

A realização desta dissertação não seria possível sem o apoio de diversas pessoas que, de diferentes formas, me permitiram ter a determinação, a força e o discernimento necessários para enfrentar os desafios propostos.

Primeiramente, quero agradecer e enaltecer o papel dos meus pais, Márcia e José Carlos, não só pelos valores e ensinamentos que me transmitiram durante 24 anos de vida, como também por todos os esforços realizados para que eu conseguisse ter as melhores experiências possíveis no decorrer da minha vida académica e profissional. O apoio incondicional, o orgulho, as memórias insubstituíveis, o carinho e a confiança permitiram-me ter força para superar todos os desafios que surgiram ao longo deste difícil processo criativo.

Ao meu irmão Lucas, agradeço por todos os momentos de diversão e relaxamento que me permitiram manter um foco saudável nos estudos e investigação. As saídas, interações, brincadeiras e jogos foram momentos importantes para desbloquear as fases mais difíceis durante este período.

Agradecer aos meus avós, Porfírio e Bernardina, pelo amor, apoio, dedicação e devoção aos seus netos. A vossa presença e carinho foram fundamentais para a realização desta dissertação e consequente realização dos meus sonhos. Ter alguém que, como vós, me quer sempre bem, fazendo, em simultâneo, inúmeros sacrifícios em prol da minha felicidade, serve como um reforço de motivação para a concretização de todos os meus objetivos.

Gostaria de deixar um caloroso agradecimento à Rita Martins, uma mulher que me acompanhou e me facultou uma panóplia de experiências incríveis. Contagiando-me com a sua felicidade genuína, ajudou-me a crescer pessoalmente, motivando-me sempre a ser a minha melhor versão. Com o teu apoio e carinho, senti-me sempre capaz de superar qualquer que fosse o obstáculo, tornando-se esta conquista mais recompensadora.

Ao Caramelo, meu fiel companheiro de 4 patas, que sempre me recebeu com a maior das euforias, mostrando-me também um estado puro de inocência e amor.

À Ana e à Bia, minhas amigas de longa data, quero agradecer pela contínua amizade e cumplicidade. É uma enorme felicidade acompanhar e ser acompanhado por pessoas que me são muito queridas há muitos anos.

Também gostaria de expressar a minha profunda gratidão à minha orientadora, Professora Cristina Aragão Seia, pelo seu zeloso acompanhamento e valiosa orientação. A

sua disponibilidade, experiência e conhecimento foram essenciais para a concretização e sucesso desta dissertação. Estou eternamente grato pelos ensinamentos e conselhos que com certeza me acompanharão por longos anos.

Por último, mas não menos importante, gostaria de agradecer a todos os meus restantes familiares e amigos que, de uma forma ou de outra, foram deixando a sua marca em todo o meu percurso. Os votos de confiança, amizade e competitividade motivaram-me a demonstrar de mim, o meu melhor.

A todos, o meu muito obrigado.

Índice

| | |
|--|-----|
| Agradecimentos | I |
| Resumo | IV |
| Palavras-chave | V |
| Abstract | VI |
| Lista de Abreviaturas | VII |
| 1. Introdução | 1 |
| 2. Inteligência Artificial..... | 5 |
| 2.1. Conceito | 5 |
| 2.2. Origem e Evolução Histórica | 9 |
| 2.3. Conceitos Inteligência Artificial | 17 |
| 2.3.1. <i>Artificial Narrow Intelligence</i> (IA Fraca) | 18 |
| 2.3.2. <i>Artificial General Intelligence</i> (IA Forte)..... | 19 |
| 2.3.3. <i>Artificial Super Intelligence</i> | 21 |
| 2.3.4. <i>Machine Learning</i> | 22 |
| 2.3.4.1. <i>Supervised Learning</i> | 24 |
| 2.3.4.2. <i>Unsupervised Learning</i> | 24 |
| 2.3.4.3. <i>Reinforcement Learning</i> | 26 |
| 2.3.5. <i>Deep Learning</i> | 28 |
| 2.4. Inteligência Artificial nos dias de hoje | 30 |
| 2.5. Futuro da Inteligência Artificial..... | 31 |
| 2.6. Inteligência Artificial e o Direito | 35 |
| 3. Responsabilidade Civil por danos causados por Sistemas de Inteligência Artificial | 38 |
| 3.1. Responsabilidade Civil – Considerações Gerais | 38 |
| 3.2. Responsabilidade Civil Extracontratual | 40 |
| 3.2.1. Responsabilidade Civil Extracontratual por factos ilícitos..... | 40 |
| 3.2.2. Responsabilidade pelo Risco | 46 |
| 3.2.2.1. A responsabilidade do comitente | 47 |
| 3.2.2.2. A responsabilidade do produtor por produtos defeituosos | 49 |
| 3.3. Responsabilidade Civil pelos danos causados por Sistemas de Inteligência Artificial | 53 |
| 3.3.1. Presunção Legal: Artigo 493.º CC..... | 60 |
| 3.3.2. Artigo 500.º CC | 64 |
| 3.3.3. Responsabilidade do produtor por produtos defeituosos..... | 65 |

| | |
|---|-----|
| 4. Abordagem Europeia à Responsabilidade da Inteligência Artificial | 76 |
| 4.1. Livro Branco sobre Inteligência Artificial..... | 76 |
| 4.2. Resolução do Parlamento Europeu de 2020 | 79 |
| 4.3. Regulamento de Inteligência Artificial | 55 |
| 4.4. Diretiva de Responsabilidade da Inteligência Artificial..... | 96 |
| 4.4.1. Proposta de Revisão da Diretiva dos Produtos Defeituosos | 97 |
| 4.4.2. Proposta de Diretiva de Responsabilidade de IA | 103 |
| 5. Conclusões Finais | 108 |
| 6. Bibliografia..... | 113 |
| 6. Webgrafia | 119 |
| 6. Jurisprudência Citada | 120 |

Resumo

Esta dissertação aborda a responsabilidade civil por danos causados por sistemas de Inteligência Artificial (IA), uma questão emergente na era tecnológica. O desenvolvimento tecnológico traz inúmeros benefícios, mas também desafios, especialmente legislativos na área da responsabilidade. Para abordar o tema central, começamos pela determinação do conceito de IA, que, devido à sua complexidade, não tem uma definição unânime, resultando assim em diversas versões conceituais. Analisaremos também o enquadramento histórico bem como alguns dos principais conceitos ligados à IA, para um melhor entendimento destes sistemas.

O problema central discutido reside nos regimes de responsabilidade civil. Questiona-se quem deve ser responsabilizado por danos causados por sistemas de IA. Serão os regimes tradicionais suficientes para proteger os lesados? Numa tentativa de responder a esta problemática, inicialmente, faremos uma análise geral dos regimes de responsabilidade civil e, em seguida, focaremos os nossos esforços nos regimes de responsabilidade extracontratual. Concluiremos que devido às características específicas dos sistemas de IA, como a opacidade, a autonomia, a conectividade, a dependência de dados, verificamos uma dificuldade na aplicação dos regimes tradicionais de responsabilidade, principalmente pela falha dos requisitos previstos na legislação atual. Esta dificuldade encaminha-nos na direção dos regimes de responsabilidade objetiva, como no caso da diretiva dos produtos defeituosos, regime que também será explorado. Ainda assim, não se encontram sanadas todas as inadequações, o que levanta a necessidade de uma intervenção legislativa no sentido de melhor acautelar estes cenários.

Por sua vez, e numa abordagem europeia, examinaremos os marcos legislativos recentes sobre IA na União Europeia, resultado dos esforços e discussões conjuntas do Parlamento Europeu e Comissão Europeia. Tem-se em vista uma harmonização das leis, buscando-se criar um regime comum que ofereça um nível de proteção adequada aos lesados, nível este similar ao oferecido pelos regimes tradicionais quando aplicados a casos que não envolvem IA. É desta forma que serão analisadas as propostas do Parlamento Europeu e da Comissão Europeia, que culminaram no Livro Branco sobre IA (2020) e no Regulamento de IA (2024), dando-se também uma especial atenção às últimas propostas em discussão: proposta de Diretiva do Parlamento Europeu e do Conselho relativa à responsabilidade decorrente dos produtos defeituosos e a proposta de Diretiva do Parlamento

Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial, mais conhecida como Diretiva de Responsabilidade da IA.

Palavras-chave

Inteligência Artificial; Responsabilidade Civil; Direito da União Europeia

Abstract

This dissertation addresses civil liability for damages caused by artificial intelligence (AI) systems, an emerging issue in the technological era. Technological development brings numerous benefits but also challenges, especially legislative ones in the field of liability. To address the central theme, we begin by defining the concept of AI, which, due to its complexity, lacks a unanimous definition, resulting in various conceptual versions. We will also analyze the historical framework and some of the key concepts related to AI for a better understanding of these systems.

The central issue discussed lies in the civil liability regimes. The question is who should be held accountable for damage caused by AI systems. Are traditional regimes sufficient to protect those harmed? In an attempt to address this issue, we will first conduct a general analysis of civil liability regimes and then focus on non-contractual liability regimes. We conclude that due to the specific characteristics of AI systems, such as opacity, autonomy, connectivity, and data dependency, there is difficulty in applying traditional liability regimes, particularly due to the failure to meet the requirements set out in current legislation. This difficulty leads us toward strict liability regimes, such as those for defective products, which will also be explored. However, not all inadequacies are fully resolved, highlighting the need for legislative intervention to better address these scenarios.

In turn, and from a European perspective, we will examine recent legislative milestones on AI within the European Union, resulting from joint efforts and discussions by the European Parliament and the European Commission. The goal is to harmonize laws, aiming to create a common regime that offers adequate protection to those harmed, similar to the protection provided by traditional regimes in cases not involving AI. This is how the proposals from the European Parliament and the European Commission will be analyzed, culminating in the White Paper on AI (2020) and the AI Regulation (2024), with special attention to the latest proposals under discussion: the proposed Directive of the European Parliament and Council on liability for defective products and the proposed Directive of the European Parliament and Council on adapting non-contractual civil liability rules to AI, also known as the AI Liability Directive.

Keywords

Artificial Intelligence; Civil Liability; European Union Law

Lista de Abreviaturas

AC. – Acórdão

AI – Artificial Intelligence

AGI – Artificial General Intelligence

ALPAC – Automatic Language Processing Advisory Committee

ANN – Artificial Neuron Network

ANI – Artificial Narrow Intelligence

ASI – Artificial Super Intelligence

BCI – Brain-Computer Interfaces

BGB – Bürgerliches Gesetzbuch (Código Civil Alemão)

CC – Código Civil

Cfr. – Conferir

CPC – Código Processo Civil

GPAN – Grupo de Peritos de Alto Nível sobre Inteligência Artificial

IA - Inteligência Artificial

IBM – International Business Machines Corporation

JRC – Joint Research Center

ML – Machine Learning

Ob. Cit. – Obra/s Citada/s

p./pp. – Página/Páginas

RIA – Regulamento Inteligência Artificial

1. Introdução

A evolução tecnológica é um dos fenômenos de maior relevância no panorama mundial há já muitos anos, profundamente ligado a um crescimento célere e impactante. Em pleno 2024, a tecnologia em si, é completamente indissociável da sociedade, já não sendo apenas um resultado do desenvolvimento humano, mas também se assumindo como um objetivo a prosseguir e a preservar.

Esta “nova” vertente tecnológica encontra-se enraizada em muitos, se não já todos, os setores fundamentais de qualquer sociedade. É perfeitamente possível fazer-se um antes e depois nas diversas áreas das atividades, observando-se claras melhorias após a referida revolução tecnológica. Em áreas como a educação, a ciência, o direito, a saúde, a indústria, a agricultura, a economia, os negócios, os transportes, a segurança, o entretenimento e a cultura, são notórias as mudanças de paradigma ao nível da eficácia, sustentabilidade, conectividade, comunicação e segurança. Pense-se, por exemplo, no impacto que a introdução dos computadores e, posteriormente, da internet tiveram no mundo, não só ao nível da simplificação e automatização de variados comportamentos, mas também a propósito da comunicação e partilha de informação. A aproximação das pessoas umas às outras, sem real diminuição do distanciamento físico, abriu verdadeiramente as portas a uma realidade totalmente diferente da verificada, a realidade da Era Digital ¹.

A inteligência artificial, apesar da sua recente popularidade, é apenas uma parte desta enorme vaga tecnológica, mas é, inegavelmente, uma das frentes mais inovadoras e ambiciosas de sempre. Através de uma correta utilização dos diferentes meios de inteligência artificial, é possível chegar-se ao próximo nível da Era Digital. O recurso, cada vez mais comum, a estes sistemas de *software* munidos de algoritmos de aprendizagem automática em áreas que já se encontravam digitalizadas, permite-nos dar um grande passo na direção de uma maior automatização e conseqüente menor supervisão ou intervenção humana, podendo eventualmente se atingir uma, já temida, substituição tecnológica.

¹ Sem prejuízo de tecnologias anteriores (que à sua maneira foram também relevantes, como a máquina de escrever, a fotografia, o rádio, a televisão, entre outros), é só com a introdução da internet que conseguimos, verdadeiramente, observar um grande salto na metodologia e funcionamento geral considerado “normal”. Plataformas *online*, ferramentas digitais que armazenam e disseminam todo um conjunto de informação, permitindo também uma superior otimização dos recursos; Indústria e agricultura com processos mais automatizados e de produção em massa com maior eficiência, segurança, quantidade e qualidade; Medidas de segurança e controlo reforçadas; Redes sociais que permitem interligar variados pontos no mundo; evolução dos meios de comunicação e introdução de novos meios como *smartphones*, e muito mais. Por outras palavras, foi com este marco histórico que tivemos o início de um processo enorme de conversão digital, uma digitalização do mundo, a velocidades diferentes.

Todas estas inovações tecnológicas estão, indubitavelmente, ligadas a vastos benefícios, não só ao nível da automação de vários procedimentos (como o controlo e gestão de mecanismos, máquinas e recursos, útil na eficiência das indústrias de produção e agricultura, ou a definição de itinerários e calendarizações, muito presentes no nosso quotidiano), como também no plano de uma assistência mais teórica, implicando para isto métodos e técnicas de raciocínio (como por exemplo, ao permitir uma maior precisão dos diagnósticos médicos ou uma melhor prevenção das doenças na área da saúde, ou ao ser utilizada como ferramenta auxiliar jurídica, como motores de pesquisa de jurisprudência, ferramentas de análise, resolução de litígios (em linha ou *online*, com a sigla RLL) ou a elaboração de documentos legais, entre outros ²). Apesar de ainda nos encontrarmos num estágio embrionário de sistemas de IA, muitos dos exemplos mencionados, já são realizáveis, com maior ou menor limitação, provando que o seu potencial e benefício são bem reais.

Contudo, enquanto se apresenta como uma tecnologia com bastantes benefícios, também comporta a sua porção de riscos e prejuízos para a sociedade bem como para os interesses públicos e direitos fundamentais dos seus cidadãos. Por se tratar de uma tecnologia que tem um impacto tão grande na sociedade, suscitam-se várias questões ao nível ético, social e jurídico, incidindo essencialmente sobre seu o desenvolvimento, utilização e segurança. Questões relacionadas com a criação de algoritmos utilizados de forma maliciosa, rompendo sistemas de segurança, gerando e disseminando desinformação, ou enviesados com critérios desiguais e discriminatórios; questões relacionadas com a rápida automatização e conseqüente substituição do ser humano em diversos postos de trabalho, levantando problemas de desemprego e, conseqüentemente, de subsistência; questões associadas a violações de direitos como a privacidade; questões sobre o desconhecimento e incorreta utilização desses sistemas e as questões visadas com a nossa temática da responsabilidade civil pelos danos que a IA pode causar.

É precisamente neste contexto mais negativo da inteligência artificial que o Direito deve de se focar, de forma que consiga elaborar as respostas mais apropriadas a problemas que vão surgindo, principalmente no âmbito da responsabilidade civil. Quem será responsável pelos danos causados por sistemas de Inteligência Artificial? Perante algoritmos programados incorretamente com ideais desiguais ou que, através da sua capacidade de recolha de dados, realizem uma colheita envenenada por diretivas injustas e discriminatórias

² Elisa Alfaia Sampaio e Paulo Jorge Gomes, “A Inteligência Artificial como auxiliar das decisões judiciais”, em “Direito e Inteligência Artificial”, coordenação de Maria Raquel Guimarães e Rute Teixeira Pedro, Almedina, 2023, p. 209.

(baseadas em preconceitos em razão do sexo, raça, língua ou religião), quem será responsável pelos danos resultantes das decisões ou conclusões desses sistemas inteligentes? Nos casos de veículos autónomos (com elevados níveis de autonomia), de eletrodomésticos, *smarthomes* e sistemas de segurança, a quem serão imputados os prejuízos resultantes do mau funcionamento ou de falhas de manutenção de segurança dos *softwares*? Será o modelo da responsabilidade do produtor pelos produtos defeituosos que coloca em circulação, adequado a solucionar estes casos? Passará a resposta pela atribuição de uma espécie de personalidade jurídica a estes entes tecnológicos, de maneira que os próprios sistemas, a própria máquina, indemnice os danos por si causados?

A dificuldade em apresentar soluções às distintas possibilidades que podem surgir deve-se à complexidade da temática em questão, pois estes sistemas de inteligência artificial apresentam um conjunto de características específicas, (como por exemplo, a conectividade, a autonomia e a dependência de dados), bem como uma pluralidade de agentes económicos envolvidos na cadeia de produção e distribuição, que tornam obsoletas, ou no mínimo, insuficientes, as disposições atuais de regimes de responsabilidade civil ³. Considerando estes aspetos especiais, bem como os tipos diferentes de aprendizagem (automática e profunda) que alargam o nível de autonomia e imprevisibilidade destes algoritmos, surge uma verdadeira incógnita na relação Direito-Inteligência Artificial.

O que torna este tema tão fundamental nas discussões ao dia de hoje, não se deve apenas ao considerável impacto positivo e negativo que estas tecnologias podem ter na sociedade, deve-se também ao facto de a inteligência artificial ser o futuro da humanidade. Apesar de ainda estarmos longe de cenários ficcionais apocalípticos, estamos a viver uma era onde há uma dependência cada vez maior das pessoas face a estas novas tecnologias, já não apenas para tarefas simples e redundantes (como a marcação e organização de itinerários, traduções imediatas, correções automáticas, recomendações de conteúdo), mas também recorrendo-se a estas tecnologias para tarefas mais complexas que exigem uma maior atenção (como por exemplo conduções automatizadas, controlo de sistemas de segurança, aconselhamento e decisões jurídicas, médicas e financeiras, entre outras).

É precisamente neste contexto de preparação e prevenção para o futuro, que as grandes instituições e governos estão a priorizar a inteligência artificial. A União Europeia não se mantém alheia a este pensamento, sendo muitos os esforços legislativos, ao longo dos

³ Relatório da Comissão ao Parlamento Europeu, ao Conselho e ao Comité Económico e Social Europeu, “Relatório sobre as implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial, da internet das coisas e da robótica”, COM (2020) 64 final, Bruxelas, 2020, p.2.

anos, na tentativa de harmonizar regras para a criação e manutenção de sistemas inteligentes mais seguros e confiáveis. Estes são objetivos que continuarão na ordem do dia da União, tal como a presidente da Comissão Europeia, Ursula von der Leyen fez questão de reconhecer nas suas mais recentes orientações políticas para a próxima Comissão Europeia 2024-2029: “Devemos agora concentrar os nossos esforços em tornarmo-nos um líder global em inovação em IA”⁴.

⁴ Ursula von der Leyen, “*Europe’s Choice: Political Guidelines for the next European Commission*” 2024-2029, p.10.

2. Inteligência Artificial

2.1. Conceito

Em primeiro lugar e, antes de qualquer avanço, apresentamos uma definição de uma das principais temáticas desta dissertação: O que é a Inteligência Artificial? Este é um tema relativamente recente, não só no que diz respeito à presença desta nova tecnologia nas mais variadas áreas como a medicina, o direito, a economia ou a educação, entre outras, mas também na própria área da ciência da computação que se dedica ao estudo e desenvolvimento dos sistemas informáticos.

A definição de Inteligência Artificial (IA) não se afigura tarefa simples, aliás esta dificuldade está demonstrada na inexistência de uma noção única e consensual, existindo, assim, várias visões daquilo que pode e deve de ser considerado Inteligência Artificial. É importante referir também o elevado nível de complexidade envolvido nesta matéria, tanto mais que se encontra numa área tão versátil como a da ciência da computação.

Uma primeira abordagem que podemos fazer a uma espécie de definição poderá ser através do método tradicional do recurso ao dicionário português. Assim, Inteligência Artificial é vista como o “*ramo da ciência da computação que estuda o desenvolvimento de sistemas computacionais com base no conhecimento sobre a inteligência humana*”⁵. Numa lógica comparativa e recorrendo ao dicionário de Cambridge, é definida como “o estudo de como produzir máquinas que apresentam as mesmas qualidades detidas pela mente humana, tal como a habilidade de compreender linguagem, reconhecer imagens, resolver problemas e aprender”⁶.

Historicamente, foram vários os cientistas computacionais que fizeram grandes avanços no campo da Inteligência Artificial, sendo um deles John McCarthy a quem se deve a utilização, pela primeira vez, do termo “*Artificial Intelligence*” (AI). McCarthy definiu IA como “*the science and engineering of making intelligent machines especially intelligent computer programs*”⁷. Outro grande cientista, Alan Turing, foi considerado o “*father of Computer Science*”, principalmente pelos seus trabalhos com o desenvolvimento do “Teste

⁵ De acordo com o Dicionário Priberam Online de Português, disponível na internet em <https://dicionario.priberam.org/>.

⁶ Traduzido do Cambridge Dictionary, disponível na internet em <https://dictionary.cambridge.org/>;

⁷ John McCarthy, “*What is Artificial Intelligence?*”, Computer Science Department, Stanford University, 2007, p.2 (disponível em <https://www-formal.stanford.edu/jmc/>); Também Christopher Manning, “*Artificial Intelligence Definitions*”, Stanford University Human-Centered Artificial Intelligence, 2020, disponível na internet em <https://hai.stanford.edu/>.

de Turing”. Alan Turing não define expressamente Inteligência Artificial, mas deixa clara a ideia de a máquina poder imitar a mente humana, enquanto procura a resposta à questão “*Can Machines think?*”. Uma Inteligência Artificial parece ser assim, para o cientista computacional, uma máquina com capacidade para imitar a mente humana, imitar o comportamento inteligente do ser humano ⁸.

Uma outra abordagem mais complexa ao conceito de Inteligência Artificial passa pela sua desconstrução em dois momentos. Num primeiro momento, é pertinente definir Inteligência Humana pois só assim saberemos o que será recriado artificialmente. Ora, inteligência é considerada o “*conjunto de todas as faculdades intelectuais como a memória, a imaginação, o juízo, o raciocínio, a abstração e a conceção*” ⁹. Perfilamos assim os esforços realizados por Howard Gardner com a sua “Teoria das Inteligências Múltiplas” onde defendeu uma visão multifacetada da mente, ou seja, para o autor, a Inteligência Humana é bastante diversificada manifestando-se numa vasta panóplia de áreas como a Inteligência Linguística, a Inteligência Musical, a Inteligência Lógico-Matemática, a Inteligência Espacial, a Inteligência Corporal-Cinestésica, a Inteligência Intrapessoal e Interpessoal, a Inteligência Naturalista e Existencial e por fim a Inteligência Criativa ¹⁰. É precisamente devido a esta grande diversidade e complexidade apresentada pela inteligência humana, que se afigura difícil a sua recriação a um nível artificial ou robótico, levando muitos autores a concluir como prematuro os cenários apocalípticos onde as máquinas, dotadas com avançados Sistemas de IA, dominam o mundo ¹¹. Por fim, o segundo momento é a artificialidade da inteligência em causa. “Artificial” é aquilo que não é natural, algo dissimulado, postíço, geralmente feito pelo ser humano na tentativa de copiar algo natural ¹². Portanto, esta visão conclui que Inteligência Artificial é a recriação, pelo homem, de algo que apresente as mesmas habilidades cognitivas que a mente humana, ou pelo menos um nível semelhante.

Richard E. Susskind, refere que IA foca-se no *design*, desenvolvimento e implementação de sistemas computacionais que conseguem realizar tarefas e solucionar problemas que normalmente exigem um determinado nível de inteligência humana. O autor

⁸ Alan M. Turing, “Computing Machinery and Intelligence”, in *MIND, A Quarterly Review of Psychology and Philosophy*, vol. LIX, n.º 236, 1950, pp.433.

⁹ Definição de acordo com o Dicionário Priberam Online de Português, disponível na internet em <https://dicionario.priberam.org/>.

¹⁰ Howard Gardner, “*Inteligências Múltiplas: a teoria na prática*”, Tradução por Maria Adriana Veríssimo Veronese, Artmed Editora, Porto Alegre, 1995, pp.13-16.

¹¹ Ralf T. Kreutzer, Marie Sirrenberg, “*Understanding Artificial Intelligence. Fundamentals, Use Cases and Methods for a Corporate AI Journey*”, Springer, Berlim, 2020, p.3.

¹² Interpretação conjunta dos Dicionários de Cambridge e Priberam Online de Português.

dá como exemplo sistemas que foram programados para compreender a linguagem e traduzi-la imediatamente de um idioma para outro, ou sistemas que reconhecem objetos e imagens no mundo físico ¹³.

Podemos ainda constatar que cada um de nós pode referir-se a coisas diferentes quando utilizamos a terminologia “Inteligência Artificial”, podendo esta ser vista como um campo científico, uma tecnologia ou método, ou como aplicações concretas de Sistemas de IA ¹⁴. Deve, portanto, ser visto o seu conceito como multifacetado, não só dada a sua complexidade, mas também devido à sua grande variabilidade.

Apesar das diversas visões do conceito de Inteligência Artificial, a recriação da mente humana parece ser um elemento em comum, estando a ser utilizada quase sempre como critério de comparação ou como padrão mínimo a atingir. Ainda assim, há quem defenda a ideia de que, embora útil, a inteligência humana é um critério de uso limitado pois já existem sistemas autónomos que demonstram *performances* super-humanas face a determinadas atividades, como traduções ou padronização em videojogos, sistemas estes que não funcionam da mesma maneira que a mente humana, nem exibem qualquer tipo de consciência na realização das ditas tarefas ^{15 16}.

Uma definição mais vaga e simplista encontra-se no entendimento de IA como uma “*disciplina que tem por objetivo o estudo e construção de entidades artificiais com capacidades cognitivas semelhantes às dos seres humanos.*”. Esta simplicidade é propositada, como aludida por Costa e Simões, e justificada com os “*diferentes pontos de*

¹³ Richard E. Susskind, “Artificial Intelligence, Expert Systems and Law”, in *The Denning Law Journal*, vol.5, nº.1, 1990, p.105.

¹⁴ Pascal D. König, Tobias D. Krafft, Wolfgang Schulz, Katharina A. Zweig, “Essence of AI: What is AI?”, in “*The Cambridge Handbook of Artificial Intelligence. Global Perspectives on Law and Ethics*”, edited by Larry A. DiMatteo, Cristina Poncibò, Michel Cannarsa, Cambridge University Press, Cambridge, 2022, p.18.

¹⁵ Pascal D. König, Tobias D. Krafft, Wolfgang Schulz, Katharina A. Zweig, *Ob. cit.*, p.23.

¹⁶ Ao dia de hoje são vários os sistemas que nos são disponibilizados das mais diversas formas (através de computadores, telemóveis, automóveis, entre outros) e que não só automatizam a grande maioria dos atos considerados fúteis, como também simplificam um vasto conjunto de tarefas. São exemplos disto as assistentes virtuais como a *Siri*, *Google Assistance*, *Cortana*, *Alexa* que por meio de comandos de voz, auxiliam diversas tarefas substituindo o tradicional toque, que por si só já era bem simples.

Também aqui se incluem os recém populares *Chatbots*, como *ChatGPT*, *Gemini*, *LaMDA*, *Replika* entre muitos outros, que se categorizam como modelos de linguagem ou Inteligências Artificiais Conversacionais (*Chatbots*), conseguindo quase que instantaneamente concretizar qualquer pedido feito pelo utilizador, desde que dentro da sua, ainda restrita, base de dados. Estes são apenas alguns exemplos de sistemas que já estão entre nós há algum tempo, ou que acabaram de ser introduzidos e que, apesar de ainda apresentarem algumas fortes limitações, já demonstram uma capacidade superior para determinadas tarefas, pelo menos em tempo de reação e concretização quando comparados com a própria mente humana.

vista e distintas fontes de inspiração” que levam alguém a tentar definir Inteligência Artificial ¹⁷.

Tal como é possível observar, não existe uma linha de pensamento seguida unanimemente por todos. Ora, esta incerteza face à definição de Inteligência Artificial é bastante prejudicial de um ponto de vista regulador pois, se não houver um entendimento no mínimo semelhante da temática em questão, as respostas do sistema jurídico não serão as mais adequadas. Foi precisamente neste sentido que foi criado pela Comissão Europeia, em 2018, o GPAN – Grupo de Peritos de Alto Nível sobre a Inteligência Artificial – (ou HLEG – *High Level Experts Group on Artificial Intelligence*), com o propósito de prestar aconselhamento sobre a sua estratégia de IA. Vários foram os relatórios reunidos por este grupo de estudos, tendo mesmo sido apurada uma definição operacional e atualizada de Inteligência Artificial: *“Os sistemas de inteligência artificial (IA) são sistemas de software (e eventualmente também de hardware) concebidos por seres humanos, que, tendo recebido um objetivo complexo, atuam na dimensão física ou digital percecionando o seu ambiente mediante a aquisição de dados, interpretando os dados estruturados ou não estruturados recolhidos, raciocinando sobre o conhecimento ou processando as informações resultantes desses dados e decidindo as melhores ações a adotar para atingir o objetivo estabelecido. Os sistemas de IA podem utilizar regras simbólicas ou aprender um modelo numérico, bem como adaptar o seu comportamento mediante uma análise do modo como o ambiente foi afetado pelas suas ações anteriores.*”¹⁸.

Sobre a definição apresentada pelo GPAN, Serena Quattrococo ¹⁹, considera-a ser altamente específica e detalhada, originando uma dificuldade de aplicação, principalmente num domínio não científico, estando por isto mais de acordo com a noção alternativa apresentada pelo JRC (*Joint Research Center*): *“IA é um termo genérico que se refere a qualquer máquina ou algoritmo capaz de observar o seu ambiente, aprender e, com base no conhecimento e experiência adquiridos, tomar ações inteligentes ou propor decisões. Hoje, há diversas tecnologias que se enquadram nesta ampla definição de IA. Agora as técnicas ML (Machine Learning) são as mais utilizadas”* ²⁰.

¹⁷ Ernesto Costa, Anabela Simões, *“Inteligência Artificial: Fundamentos e aplicações”*, 2ª edição revista e aumentada, FCA, Lisboa, 2008, p.3.

¹⁸ Grupo de Peritos de Alto Nível sobre a Inteligência Artificial, *“Uma definição de IA: Principais Capacidades e disciplinas científicas”*, Comissão Europeia, disponível na internet em <https://digital-strategy.ec.europa.eu/pt/policies/expert-group-ai>.

¹⁹ Serena Quattrococo, *“Artificial Intelligence, Computational Modelling and Criminal Proceedings: A Framework for a European Legal Discussion”*, Springer, 2020, pp.7 e 8.

²⁰ *“AI is a generic term that refers to any machine or algorithm that is capable of observing its environment, learning, and based on the knowledge and experience gained, taking intelligent action or proposing decisions.*

Concluimos, assim, que a definição de Inteligência Artificial é algo verdadeiramente desafiador, até mesmo para os grandes cientistas computacionais. A inexistência de um conceito unânime obviamente não resulta da falta de esforços, sendo esta uma das temáticas mais debatidas recentemente. Deve-se, maioritariamente, ao facto de a inteligência ter-se revelado uma “*propriedade difícil de caracterizar e de reproduzir*”²¹. Alcançar uma réplica num suporte digital, da “*versatilidade, adaptabilidade e robustez que o cérebro humano exhibe ao processar informação*”²² não é de todo tarefa fácil e, pelo que até ao momento se conhece, tal feito ainda não foi verdadeiramente alcançado. Imitar a liberdade, criatividade e flexibilidade demonstrada pela mente humana será das coisas mais difíceis de se reproduzir artificialmente²³. Se é recomendável ou não dissipar a linha que separa a humanidade da máquina, é todo um outro debate.

Resta, portanto, referir que de entre todas as definições apresentadas e caminhos possíveis, aquela que nos parece mais compatível com o principal rumo deste projeto é uma mistura híbrida entre os conceitos apresentados tanto pelo GPAN como pelo JRC, ou seja, olhar para Inteligência artificial como um *Software*, independentemente do *Hardware* em que esteja incorporado, composto por um algoritmo com capacidade de observar o seu ambiente, recolher e interpretar dados, aprender e, baseando-se no conhecimento e experiência obtida com interações passadas, tomar decisões inteligentes e propor diversas vias de ação.

2.2. Origem e Evolução Histórica

Nos dias de hoje, a temática da Inteligência Artificial ainda é algo relativamente recente e começa cada vez mais a fazer parte dos grandes debates mundiais. Ainda assim, já detém uma história de certa forma densa e longa.

Nos seus primórdios podemos recuar tão longe como à Grécia antiga (384-322 a.C.) onde desde muito cedo se começou a discutir com Aristóteles, os silogismos, que eram a

There are many different technologies that fall under this broad AI Definitions. Now, ML (Machine Learning) techniques are the most widely used, Joint Research Center Technical Reports, “*AI Watch: Defining Artificial Intelligence*”, Luxemburgo, 2020, p.9.

²¹ Arlindo Oliveira, “*Inteligência Artificial*”, Fundação Francisco Manuel dos Santos, 2019, p.7.

²² *Ibidem*.

²³ O “*processamento de informação que permite a um animal, ainda que simples, reconhecer um predador, identificar uma fonte de alimento ou reagir a um perigo exterior, é extremamente complexo, exige inteligência e não é fácil de reproduzir num sistema artificial.*” O cérebro humano tem muitas “*funções básicas*” e “*capacidades inatas*” que com alguma aprendizagem e treino permitem a realização de tarefas ainda mais complexas, isto é o que se pode perceber como exibir inteligência. Artificialmente podemos nos aproximar de um nível muito semelhante a esse, mas a recriação da consciência, daquilo que nos faz humanos, ainda estará num futuro distante. Arlindo Oliveira, *Ob. cit.*, pp.25-27.

possibilidade de se formalizar o pensamento por meio de um conjunto de regras ²⁴, o que deu origem a importantes desenvolvimentos na área da lógica, fundamental para a Inteligência Artificial.

Mais tarde, no decorrer do século XVII, é possível observar também um aumento do interesse pela automatização, justificado pelos diversos desenvolvimentos em variadas áreas como a filosofia, a matemática, a astronomia e a ciência. René Descartes foi um dos primeiros que questionou o que aconteceria se alguém criasse máquinas sofisticadas de forma humana que conseguissem andar como nós. Neste pensamento, o filósofo defendeu que as máquinas não poderiam possuir inteligência humana, porque apesar de poderem reagir a estímulos, elas não conseguiam reagir corretamente a todas as situações da mesma maneira complexa que até o mais simples ser humano conseguia. Mais, disse também que, lhes faltavam capacidades de raciocínio geral, originando ações limitadas e baseadas nas suas programações, sendo que o raciocínio é um instrumento universal que pode ser usado em qualquer situação ²⁵.

Ainda dentro da mesma época, também Thomas Hobbes prestou bons contributos ao permitir um melhor entendimento de inteligência, referindo na sua obra “Leviatã”, que o “*pensamento humano resulta, simplesmente, da manipulação de símbolos*”, apresentando uma “*visão mecanicista de inteligência*”, ou seja, que podia ser reproduzida usando um “processo mecanizado” ²⁶.

Muitas outras mentes foram desbravando esta área com pequenos passos que rapidamente se tornaram enormes pela grande revolução que provocaram.

Em 1642, Blaise Pascal construiu a “*Pascaline*”, ou Máquina de Pascal, uma máquina calculadora que fazia somas e subtrações usando rodas dentadas ²⁷.

Inspirado por esta criação, Gottfried Wilhelm Leibniz, em 1673, ultrapassou Pascal com a introdução de uma nova máquina que não só somava e subtraía, mas também conseguia multiplicar e calcular raízes ²⁸.

²⁴ Ernesto Costa, Anabela Simões, *Ob. cit.*, p.22.

²⁵ Masahiro Morioka, “Descartes and Artificial Intelligence”, *Journal of Philosophy*, vol.13, n.1, Janeiro 2023, pp.1-4.

²⁶ Arlindo Oliveira, *Ob. cit.*, p.37.

²⁷ *Ibidem*, p.38.

²⁸ Stuart Russel, Peter Norvig, “*Artificial Intelligence: A Modern Approach*”, Fourth Edition, Pearson Education Limited, 2021, p.6.

Também Charles Babbage projetou duas máquinas: a máquina diferencial e a máquina analítica ²⁹. A primeira era uma máquina de calcular automática usada para calcular tabelas matemáticas de maneira mais automatizada, quando comparada com os seus precedentes de Pascal e Leibniz. A segunda, muito mais ambiciosa, permitia executar sequências arbitrárias de instruções, contendo memória, programas armazenados e um processo mecânico que realizava operações aritméticas ou lógicas ³⁰.

Alan Turing foi outra das grandes mentes que marcaram o campo da Inteligência Artificial. Idealizou, em 1936, que “*qualquer computador com uma memória suficientemente grande, que manipule símbolos e que satisfaça algumas condições simples, consegue fazer os mesmos cálculos e obter os mesmos resultados que qualquer outro computador*” ³¹. Esta sua teoria ficou conhecida como “*The Universal Turing Machine*” (Máquina Universal de Turing), a “*Turing machine que conseguia realizar o trabalho de qualquer outra Turing machine, desde que a descrição da máquina, a ser imitada, fosse colocada numa fita para aquela ler*” ³².

Mais tarde, Alonzo Church publicou um mecanismo de cálculo, o chamado cálculo de “LAMBDA”, que era para Turing, uma alternativa ao seu trabalho, “*doing the same things in a different way*” ³³, dando a origem à conhecida Tese de Church-Turing que dita que “*qualquer resultado que possa ser efetivamente calculado pode ser calculado por uma máquina de Turing, ou por qualquer outro computador que manipule símbolos e disponha de memória suficiente.*” ³⁴.

Olhando juntamente para a visão mecanicista de Hobbes, onde a inteligência humana é vista como uma manipulação de símbolos, e para a tese de Church-Turing, que refere que “*todas as máquinas que manipulam símbolos são equivalentes entre si*”, podemos concluir que “*um computador (...) deverá ser capaz de efetuar a mesma manipulação de símbolos que um cérebro humano*” ³⁵. Esta foi uma questão claramente explorada e desenvolvida por Alan Turing nos trabalhos seguintes, bastante conceituados no mundo científico.

²⁹ Nenhuma delas foram concluídas com sucesso por Babbage, pelo menos não totalmente. Só em 1991 é que foi construída a máquina diferencial, estando agora exposta no Museu da Ciência em Londres. (Stuart Russel, Peter Norvig, *Ob. cit.*, p.14).

³⁰ Arlindo Oliveira, *Ob. cit.*, p.38/39 e Stuart Russel, Peter Norvig, *Ob. cit.*, p.14.

³¹ Arlindo Oliveira, *Ob. cit.*, p.40.

³² “*Turing machine that could perform the work of any other Turing machine, provided a description of the machine to be imitated were placed upon the tape for it to read*”, Andrew Hodges, “Alan Turing and the Turing Machine”, in *The Universal Turing Machine: A Half-Century Survey*, edited by Rolf Herken, Springer-Verlag, Second Edition, New York, 1995, p.4 e 5.

³³ *Ibidem*.

³⁴ Arlindo Oliveira, *Ob. cit.*, p.41.

³⁵ *Ibidem*, p.42.

Em 1943, McCulloch e Pitts estabeleceram as primeiras bases para as redes neurais e, propondo um modelo de neurónios artificiais que respondiam aos estímulos recebidos pelos neurónios vizinhos, demonstraram que conseguiam realizar operações Booleanas (*Boolean Logical Operations*), que são palavras simples, baseadas em valores de verdadeiro ou falso, tais como *AND*, *OR* ou *NOT* ³⁶. Desta maneira, e com estas operações lógicas, idealizaram que qualquer função computável podia ser calculada por uma rede de neurónios interligada entre si ³⁷. Apesar da enorme descoberta, a implementação deste método ficaria limitada pelo ainda fraco *hardware* da época, sendo apenas em 1951 que Marvin Minsky realizou a implementação computacional de uma rede neuronal ³⁸. Mais tarde, em 1958, o trabalho foi continuado por Frank Rosenblatt que criou o “*Perceptron*”, que consistia numa camada dupla de neurónios artificiais que geravam respostas de sim/não ³⁹. Estes foram contributos fundamentais para o desenvolvimento do “*Machine Learning*” e também do “*Deep Learning*”.

Os trabalhos nas redes neurais permitiram chegar à conclusão de que seria possível, através de “*machine learning*”, a programação de um computador para aprender e assim evoluir, conseguindo fazer mais do que as funções que lhe tinham sido inicialmente introduzidas, isto porque as redes neurais conseguiam adaptar-se internamente com base na informação que era recolhida e processada. Uma boa demonstração dessa capacidade foi realizada por Arthur Samuel, com o seu programa “*Checkers*” (1959) ⁴⁰, um programa de computador que jogava damas e estava dotado de um sistema que aprendeu a escolher as melhores jogadas possíveis, conseguindo mesmo vencer alguns jogadores avançados no jogo de tabuleiro, tudo com recurso ao “*machine learning*” ⁴¹. O autor refere ainda que os

³⁶ IBM Documentation, “*Logical (Boolean) Operators*”, disponível na internet em <https://www.ibm.com/docs/en/zos/3.1.0?topic=expressions-logical-boolean-operators>.

³⁷ Stuart Russel, Peter Norvig, *Ob. cit.*, p.16.

³⁸ A sua criação ficou conhecida como “*Snarc*” (*Stochastic Neural Analog Reinforcement Calculator*), ou também conhecido como Calculadora Neural-Análoga Estocástica de Reforço. Um dos primeiros sistemas de redes neurais artificiais, que estimulava uma rede de 40 neurónios. Stuart Russel, Peter Norvig, *Ob. cit.*, p.16.

³⁹ Pascal D. König, Tobias D. Krafft, Wolfgang Schulz, Katharina A. Zweig, *Ob. cit.*, p.20.

⁴⁰ “*Enough work has been done to verify the fact that a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program. (...) it can learn to do this in a remarkably short period of time (8 or 19 hours of machine-plying time) when given only the rules of the game, a sense of direction, and a redundant and incomplete list of parameters (...)*.” Arthur L. Samuel, “Some studies in Machine Learning Using the Game of Checkers”, in *IBM Journal of Research and Development*, vol.3, n.3, july 1959, pp.210-229.

⁴¹ Pascal D. König, Tobias D. Krafft, Wolfgang Schulz, Katharina A. Zweig, *Ob. cit.*, p.20; Arthur L. Samuel, *Ob. cit.*, p.212.

avanços verificados com a experiência seriam claramente “*applicable to many other situations*”⁴².

De entre as mais importantes investigações protagonizadas por Turing ao longo dos anos, temos de dar realce ao “*Turing Test*” (Teste de Turing) a que se refere o artigo da revista académica *Mind*, publicada em 1950, pela Universidade de Oxford, com o título “*Computing Machinery and Intelligence*”⁴³. Neste artigo, onde o autor discute a questão “*Can machines think?*”, é proposto um teste empírico que permite apreciar a capacidade de um computador exibir comportamentos inteligentes equivalentes aos dos humanos ou, no mínimo, indistinguíveis. O teste é uma espécie de “jogo de imitação” e consiste em três jogadores: **A** – uma mulher, **B** – um homem e **C** – um interrogador. O interrogador encontra-se num quarto, separado dos outros dois intervenientes, e o seu objetivo é determinar quem, entre **A** e **B**, é a mulher e quem é o homem. **A**, por sua vez, tem de proteger a sua identidade e tentar fazer com que **C**, interrogador, falhe no seu objetivo de identificação, enquanto **B**, tem de ajudar **C** e orientá-lo para que o seu palpite seja o correto. Claro que todo o jogo será feito através de um sistema de comunicação escrito ou outro que não revele a identidade de **A** e **B**, para que a decisão de **C** seja tomada com base nas interações que terá com cada jogador. De seguida, Turing questionou o que aconteceria se uma máquina fizesse parte deste jogo no lugar de **A** ou **B**. “*Will the interrogator decide wrongly as often when the game is played like this, as he does when the game is played between a man and a woman?*”⁴⁴, originando assim o teste de Turing, que continua atual, mas também comporta algumas objeções. Um computador, ou sistema de inteligência artificial, passará no teste se **C**, interrogador humano, depois das respetivas interações, não conseguir determinar se está a comunicar com uma máquina, ou seja, não conseguir distinguir a máquina do ser humano, puramente com base num “diálogo”⁴⁵.

Para a realização deste teste, Russel e Norvig mencionam a necessidade de o computador possuir certas capacidades como o processamento natural de linguagem para se comunicar, raciocínio automático para chegar a novas respostas com os dados obtidos, conhecimento e *machine learning* (aprendizagem automática) de forma a adaptar-se às novas circunstâncias. Isto numa primeira fase sem contacto físico, pois se o que estiver em causa for o *total Turing Test* então aí já será incluído um sinal de vídeo e interações físicas

⁴² Arthur L. Samuel, *Ob. cit.*, p.210.

⁴³ Alan M. Turing, “Computing Machinery and Intelligence”, in *MIND, A Quarterly Review of Psychology and Philosophy*, vol. LIX, n.236, 1950.

⁴⁴ Alan Turing, *Ob. cit.*, p.434.

⁴⁵ Stuart Russel, Peter Norvig, *Ob. cit.*, p.2.

com o interrogador, o que leva a que o computador necessite também de visão para compreender os objetos e de robótica para os poder manipular e mover ⁴⁶.

É apenas em 1956 que o termo “*Artificial Intelligence*” é cunhado por John McCarthy numa conferência em Dartmouth ⁴⁷, que organizou e onde pretendeu reunir outros investigadores com interesse na “*Automata Theory*”. Na proposta feita por McCarthy ele afirma que o objetivo dessa colaboração era o de se proceder à criação de uma máquina que conseguisse simular a inteligência humana: “*We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The Study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer*” ⁴⁸.

Os primeiros pares de anos de investigação nesta temática inovadora da Inteligência Artificial tiveram bastantes contributos teóricos e práticos de sucesso, sendo praticamente regulares as surpresas causadas por todas estas brilhantes mentes.

Foi também em 1956, com Allen Newell e Herbert Simon, que se atingiu o primeiro Sistema de Inteligência Artificial, denominado de “*Logic Theorist*” (LT) ⁴⁹. Tratava-se de um programa de computador que, segundo os seus criadores, tinha a capacidade de pensar de uma maneira não numérica, ultrapassando assim o problema persistente da “mente-

⁴⁶ *Ibidem*, p.3.

⁴⁷ Reuniu o que muitos chamam de “*The Founding Fathers of AI*”, incluindo assim John McCarthy, Claude Shannon, Marvin L. Minsky, Nathaniel Rochester, Allen Newell, Arthur Samuel, Trenchard More, Ray Solomonoff, Herbert Simon e Oliver Selfridge.

⁴⁸ “Propomos que um estudo de inteligência artificial, com duração de 2 meses e participação de 10 pessoas, seja realizado durante o verão de 1956 na universidade de Dartmouth em Hanover, New Hampshire. Este estudo parte da conjectura de que cada aspeto de aprendizagem ou qualquer outra característica de inteligência pode, à partida, ser descrito com tal precisão que uma máquina pode ser construída para simulá-la. Será feita uma tentativa para descobrir como fazer com que as máquinas usem linguagem, formem abstrações e conceitos, resolvam tipos de problemas que se encontram agora reservados para seres humanos, e evoluam. Acreditamos que podem ser feitos avanços significativos em um ou mais problemas se o grupo de estudo for cuidadosamente selecionado, para trabalharem juntos durante o verão.”, John McCarthy, Marvin Minsky, Nathaniel Rochester, Claude Shannon, “*A Proposal for the Dartmouth summer research project on artificial intelligence*”, Department of Mathematics, Dartmouth College, Hanover, 1955.

⁴⁹ Seis meses depois da Conferência de Dartmouth, Herbert Simon anuncia aos seus alunos que juntamente com Allen Newell elaboraram no papel uma “máquina pensante”, que quando concretizada na prática se denominaria de “*Logic Theorist*”. “*Over the Christmas, Al and I invented a thinking machine*”. Thomas Haigh, “Historical Reflections – Conjoined Twins: Artificial Intelligence and the Invention of Computer Science”, in *Communications of the ACM*, V.66, n.6, junho 2023, p.34.

corpo” e aproximando-se de um certo nível de raciocínio. Este programa conseguiu provar com sucesso a grande maioria dos teoremas do livro “*Principia Mathematica*” de Bertrand Russell e Alfred North Whitehead ⁵⁰. Por outras palavras, o programa tratava o raciocínio como uma árvore, onde todas as ramificações representavam possíveis operações, havendo ainda um sistema que “podava” as rotas poucos promissoras, aplicando regras heurísticas de forma a permitir uma gestão mais eficaz ⁵¹.

Mais tarde, Newell e Simon foram ainda mais longe com a criação do “*General Problem Solver*” ou GPS (1957-1960), um programa que pretendia imitar um ser humano na resolução de variados problemas, sendo assim um dos primeiros programas a atingir uma abordagem mais humana de pensamento (“*thinking humanly*”) o que demonstrou que um computador, com um algoritmo genérico, podia resolver problemas de natureza geral ⁵².

Um outro exemplo de sucesso foi a linguagem de programação de alto nível, “LISP” ⁵³ criada por John McCarthy, em 1958. Foi adotada como a principal linguagem de investigação em IA, facilitando a programação de sistemas de raciocínio baseados em lógica. McCarthy desenhou este programa, baseado no cálculo de “LAMBDA” de Alonzo Church, para utilizar o conhecimento geral do mundo e assim procurar soluções para diversos problemas ⁵⁴.

Não restam dúvidas que os anos 50 e 60 foram uma espécie de “Era Dourada” para a Inteligência Artificial não só pelo crescente interesse do mundo científico nestas novas tecnologias, mas também devido aos inúmeros avanços e descobertas concretizadas que, ainda assim, foram fortemente limitadas pelo estado do *hardware* existente, bem como pela complexidade que esta matéria comporta. O mesmo não se pode dizer dos anos que se seguiram.

A partir de 1970 e até meados de 1980, dá-se o período que ficou conhecido como o primeiro “AI Winter”, ou seja, numa tradução literal, o Inverno da IA. É um período onde a progressão da Inteligência Artificial ficou marcada por um forte desinvestimento na investigação. Maus resultados e promessas falhadas ⁵⁵, maioritariamente devido à fraca

⁵⁰ Stuart Russel, Peter Norvig, *Ob. cit.*, p.17/18.

⁵¹ Thomas Haigh, *Ob. cit.*, p.34.

⁵² Stuart Russel, Peter Norvig, *Ob. cit.*, p.18.

⁵³ LISP, é um acrónimo derivado de “*List Processing*”. John McCarthy, “*History of Lisp*”, Artificial Intelligence Laboratory, Stanford University, february, 1979.

⁵⁴ Stuart Russel, Peter Norvig, *Ob. cit.*, p.19.

⁵⁵ A decisão de desinvestimento por parte do governo americano e britânico foi desencadeada pela falta de resultados prometidos pelos departamentos de investigação de Inteligência Artificial. No caso do governo americano, umas das principais razões que justificou o interesse na pesquisa de IA foi a promessa da criação de uma máquina de tradução automática e instantânea que permitisse uma melhor compreensão do russo, dadas

capacidade do *hardware*, levaram a que os governos e exércitos se sentissem relutantes nos contínuos financiamentos dos projetos militares e científicos que envolvessem qualquer tipo de Inteligência Artificial ⁵⁶.

Apesar da fase negativa que assolou a investigação da Inteligência Artificial, conseguiu verificar-se algum desenvolvimento protagonizado por aqueles que persistiram, mesmo com um desânimo geral e uma tremenda falta de financiamento. Uma das consequências deste período foi a divisão da Inteligência Artificial em diferentes áreas como a representação de conhecimento, meios de busca, jogos, prova de teoremas, planeamento, probabilística, raciocínio, aprendizagem, processamento natural de linguagem, visão, robótica, sistemas especializados, redes neuronais, algoritmos genéticos, entre outros ⁵⁷. Assim conseguiu-se desenvolver sistemas de IA que, baseando-se em diversas formas de manipulação de símbolos, permitiram criar muitos dos algoritmos que são executados pelos computadores atuais, como por exemplo os algoritmos que permitem fazer uma otimização de horários de comboios, aviões e outros meios de transporte ⁵⁸.

É com o surgimento dos “*Expert Systems*” que começamos a ver uma mudança de panorama. São sistemas que procuram apurar respostas a vários problemas recorrendo a conhecimento especializado em certas áreas, normalmente detido por um ser humano especialista na matéria ⁵⁹. Ou seja, há uma maior eficácia destes programas pois estão limitados a uma área de especialidade, obtendo-se assim uma maior concentração do “raciocínio”, ao contrário dos esforços passados focados na criação de sistemas de

as tensões vividas durante a Guerra Fria. Sendo assim um dos gatilhos que serviu para a retirada do financiamento foi o relatório negativo do ALPAC (Automatic Language Processing Advisory Committee) em 1966, que conclui que o *Georgetown-IBM Experiment* (uma das primeiras tentativas de criação de uma máquina de tradução) não era uma máquina útil de tradução devido às suas limitações, terminando o relatório por referir que tal marco não se atingiria tão cedo. (“*We do not have useful machine translation and there is no immediate or predictable prospect of useful machine translation*”. John Hutchins, “ALPAC: The (In)famous report, in *MT News International*, n.º.14, junho, 1996, p.12).

Por outro lado, para o governo britânico foram as conclusões realizadas por Sir James Lighthill no seu relatório “*Artificial Intelligence: a General Survey*” (1972), que levaram a uma queda drástica do investimento. O ilustre professor de Cambridge, depois da sua exposição na temática da IA, desenvolveu um capítulo intitulado “*Past Disappointments*”, onde refere que nenhuma das experiências passadas tiveram o impacto prometido. (“*In no part of the field have the discoveries made so far produced the major impact that was then promised*”. James Lighthill, “*Lighthill Report: Artificial Intelligence: a paper symposium*”, Science Research Council, 1973, p.8/9).

⁵⁶ Thomas Haigh, *Ob. cit.*, p.37.

⁵⁷ Gheorghe Tecuci, “Artificial Intelligence”, in *WIREs Computational Statistics*, V.4, n.º.2, março/abril, 2012, pp.169/170.

⁵⁸ Arlindo Oliveira, *Ob. cit.*, p.56.

⁵⁹ Richard E. Susskind, *Ob.cit.*, p.105.

finalidades múltiplas ⁶⁰. O DENDRAL ⁶¹, em 1969, foi um dos primeiros exemplos destes sistemas. Outro exemplo de grande sucesso foi o MYCIN ⁶².

A Inteligência Artificial, como uma das áreas mais complexas das ciências, ainda passou por muitos mais altos e baixos com o passar dos anos, chegando mesmo a ocorrer um outro Inverno da IA, e até mesmo um terceiro, de acordo com alguns autores, mas apesar disso persistiu até ao século XIX, onde realmente se assistiu a uma explosão ao nível do seu progresso científico, tornando-a hoje numa das principais áreas de investigação.

2.3. Conceitos da Inteligência Artificial

Com o passar dos anos e com o conseqüente crescimento da investigação em matérias das ciências de computação, várias são as teorias desenvolvidas e, ainda mais, são as conclusões tiradas. Nada disto é anormal quando se tem em conta que o destino previsível da humanidade é cada vez mais digital e autónomo, fortemente dependente de inteligências artificiais. Um pouco como já se vê atualmente, mas que, em maior escala, ainda não se prevê que aconteça a curto prazo. Contudo, antes mesmo de olharmos para o estado atual da tecnologia e posteriormente para o seu futuro, é relevante delinear mais alguns conceitos que consideramos fundamentais.

Antes de mais, a Inteligência Artificial pode ser classificada em três estágios: “Artificial Narrow Intelligence” (ANI), “Artificial General Intelligence” (AGI) e por fim “Artificial Super Intelligence” (ASI).

Como um subcampo da Inteligência Artificial, torna-se também relevante desenvolver os conceitos relacionados com a aprendizagem, desde a automática (*Machine Learning*), à profunda (*Deep Learning*), e também a Supervisionada, Não Supervisionada e a Por Reforço.

⁶⁰ Stuart Russel, Peter Norvig, *Ob. cit.*, p.22.

⁶¹ Começou a ser desenvolvido por volta de 1965 na Universidade de Stanford por vários colaboradores, entre eles Edward Feigenbaum e Joshua Lederberg. Um sistema especialista em análise química e criado para auxiliar na introdução das informações fornecidas por um espectrômetro de massa na estrutura molecular. Stuart Russel, Peter Norvig, *Ob. cit.*, p.22/23.

⁶² Os trabalhos a envolver o MYCIN iniciaram-se na Universidade de Stanford em 1972, e este sistema especializado permitia realizar diagnósticos médicos. Com a informação recolhida conseguia suscitar hipóteses, requisitar mais testes laboratoriais e atingir diagnósticos prováveis bem como recomendar meios indicados de tratamento. (Jack Copeland, “MYCIN”, Britannica, disponível na internet em <https://www.britannica.com/technology/MYCIN>) Apresentava uma grande diferença com o DENDRAL pois, este era regido por modelos teóricos gerais da química, enquanto o MYCIN necessitou de várias entrevistas realizadas a peritos e médicos para que se pudesse formar a sua base de dados. (Stuart Russel, Peter Norvig, *Ob. cit.*, p.23).

2.3.1. Artificial Narrow Intelligence (IA Fraca)

A *Artificial Narrow Intelligence* ou também conhecida como “Weak AI” - Inteligência Artificial Fraca ou Limitada, é um IA treinada e focada em realizar tarefas específicas, singulares ou limitadas ⁶³.

John Searle, professor do Departamento de Filosofia da Universidade da Califórnia, define A ANI, (ou “*Cautious AI*” segunda as suas palavras), no seu artigo “Minds, brains and programs”, publicado na revista “*Behavioral and Brains Sciences*” em 1980, como uma ferramenta poderosa que auxilia o estudo da mente ao permitir simular hipóteses de maneira rigorosa e precisa ⁶⁴.

Denominar este tipo de IA como “fraca” pode induzir em erro, pois na verdade de fraca ela nada tem, é apenas um dos estágios iniciais quando olhamos para o enorme potencial que estes sistemas comportam, por isso é que é mais adequada a terminologia “narrow”, estreita.

É precisamente neste estágio tecnológico de IA que nos encontramos atualmente.

São os sistemas de ANI que dão vida a muitos pedaços de tecnologia do nosso quotidiano, como por exemplo: as já conhecidas assistentes virtuais, *Alexa* (Amazon), *Siri* (Apple), *Google Assistant*, *Cortana* (Microsoft), *Bixby* (Samsung), que simplificam uma panóplia de atos fúteis através de vários dispositivos, sejam eles telemóveis ou computadores; os modelos de linguagem ou *Chatbots*, *ChatGPT*, *Gemini*, *Replika*, que conseguem fornecer variados serviços desde traduções a explicações ⁶⁵; também os já comuns veículos autónomos como o Tesla e o Polestar, juntamente com outras marcas que

⁶³ IBM, “*What is Artificial Intelligence?*”, disponível na internet em: <https://www.ibm.com/topics/artificial-intelligence>.

⁶⁴ “*According to Weak AI, the principal value of the computer in the study of the mind is that it gives us a very powerful tool. For example, it enables us to formulate and test hypotheses in a more rigorous and precise fashion*”. John Searle, “Minds, brains and programs”, in *Behavioral and Brains Sciences*, vol.3, 1980, p.417.

⁶⁵ Numa dinâmica a propósito do conceito de ANI, foi realizada a seguinte questão, a dois dos mais utilizados modelos de linguagem atuais, *Gemini* (Experimental) e *ChatGPT* (3.5): “És uma Artificial Narrow Intelligence?”. As respostas obtidas foram as seguintes:

Gemini: “Sim, eu sou um ANI. Sou um grande modelo de linguagem, também conhecido como IA conversacional ou chatbot treinado para ser informativo e abrangente (...)”. *Gemini*, disponível em <https://gemini.google.com/chat>;

ChatGPT: “Sim, eu sou um exemplo de Artificial Narrow Intelligence (ANI). Minha capacidade principal é processar e gerar texto com base em padrões aprendidos durante o treinamento. Minha especialização está na interação baseada em linguagem natural, e eu não tenho capacidade de realizar tarefas fora desse escopo. (...)”. *ChatGPT*, disponível em <https://chat.openai.com/>;

Embora ambos tenham uma abordagem bastante natural e direta, é claramente notória a falta de consciência dos dois modelos de ANI, respondendo apenas dentro daquilo que o seu treinamento e programação delimitaram, sendo que, quando o utilizador ousa questionar algo para o qual não foram preparados, a resposta é quase sempre a mesma: “Eu sou um programa de computador desenvolvido para ajudar a fornecer informações e responder às suas perguntas da melhor forma possível, contudo não tenho opiniões, sentimentos ou gostos.”, reforçando assim a sua vertente “Narrow”.

começam a apresentar os seus sistemas autónomos alternativos, e, por fim, variados outros exemplos que se podem classificar como sistemas de recomendação *online* de motores de busca e de outros *websites*, sistemas de tradução automática, entre muitos outros ⁶⁶.

Portanto, este é um tipo de Inteligência Artificial com uma eficaz capacidade de realizar determinadas tarefas para as quais é programada, mas sem qualquer nível de consciência ou compreensão avançada. Um simples cumprir de regras estabelecidas pelos algoritmos e bases de dados. É o nível mais inferior desta tipologia, mas que pela qualidade e quantidade dos exemplos mencionados, constata-se que não é de todo uma Inteligência Artificial fraca, aliás, pode até mesmo superar a capacidade e eficácia humana na realização dessas tarefas específicas, mas sem o objetivo de simular a mente humana, querendo-se apenas automatizar determinados comportamentos.

2.3.2. Artificial General Intelligence (IA Forte)

Por outro lado, quando falamos em AGI, ou IA Forte ou Geral, estamos a referir-nos a uma forma teórica de IA, onde uma máquina irá apresentar uma inteligência equiparada ou superior à do próprio ser humano. Ou seja, tal como o homem, essa IA teria a capacidade de pensar, aprender, resolver problemas, planear para o futuro e de se adaptar da mesma maneira que a mente humana o permite. A tudo isto acresce a enorme complexidade e importante consciência do cérebro, que lhe permitiria sentir e expressar emoções ⁶⁷.

Para Searle, contrariamente à sua visão de “Weak AI” como uma mera ferramenta, aqui a “Strong AI” já não será apenas uma ferramenta de auxílio à mente, mas sim a própria mente. Segundo o autor, um programa de computador, programado corretamente, poderá literalmente compreender e apresentar outras capacidades cognitivas ⁶⁸.

Estamos, portanto, a falar de uma IA que apresenta capacidades cognitivas que se assemelham às dos seres humanos, e que podem até mesmo superá-las. Sendo assim, consegue não só compreender, como também aprender, adaptar-se e evoluir, demonstrando também um certo nível de criatividade e consciência só visto na humanidade.

⁶⁶ IBM, “*What is Artificial Intelligence?*”, disponível na internet em: <https://www.ibm.com/topics/artificial-intelligence>.

⁶⁷ IBM, “*What is Artificial Intelligence?*”, disponível na internet em: <https://www.ibm.com/topics/artificial-intelligence>.

⁶⁸ “*But according to Strong AI, the computer is not merely a tool in the study of the mind; rather the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have cognitive states.*” John Searle, *Ob. cit.*, p.417.

Quanto a este estágio, parece correto afirmar-se que, até ao dia de hoje e até onde o conhecimento público chega, ainda não foi possível a criação de AGI, ou de algo que se possa aproximar-lhe.

Vários são os cientistas que questionam se a AGI é realmente possível ou se não passa de um cenário fictício e que apenas terá concretização no mundo de Hollywood. Há quem acredite e aponte para um prazo mais curto, como cinco a dez anos, enquanto outros, menos otimistas, já preveem uma centena ou mais de anos ⁶⁹.

É caso para dizer que com sistemas de AGI, o Teste de Turing poderia ser ultrapassado com sucesso, pois seria possível manter-se uma conversa com uma máquina de uma maneira tão natural que tal sistema seria totalmente indistinguível de um ser humano. É por isto que esta visão ainda se apresenta atualmente como ficção ⁷⁰.

A criação de uma genuína AGI é o grande objetivo a ser alcançado pela humanidade e são vários os esforços e recursos direcionados, por várias empresas neste sentido. Um exemplo com algum sucesso é o da DeepMind, empresa que foi adquirida pela Google, em 2014, para se juntar aos seus laboratórios de pesquisa em IA. O diretor do Departamento de Inteligência Artificial da Google, Jeffrey Dean, admite que as diversas equipas à disposição da Google têm focos diferentes, pois enquanto uns desenvolvem os problemas relacionados com a perceção das máquinas, outros desenvolvem o entendimento natural da linguagem, havendo ainda espaço para a investigação e criação de novo *software* e *hardware*, que melhor capitalize os algoritmos de *Machine Learning* e das técnicas que vão surgindo no mundo científico ⁷¹.

A DeepMind ⁷² lidera a frente de pesquisa na procura de um método de criação de AGI. Grande parte da equipa tem vastos conhecimentos não só em IA como também em neurociência, pois defendem que esta, influencia aquela ⁷³. Parece claro que a melhor abordagem para a recriação artificial da mente humana tem de resultar da inspiração obtida através do funcionamento interno do cérebro ⁷⁴.

⁶⁹ Martin Ford, “*Rule of the Robots. How Artificial Intelligence Will Transform Everything*”, tradução por Luís Santos, “*O Futuro da Inteligência Artificial. Uma transformação que já está a acontecer*”, Bertrand Editora, 2022, p.150.

⁷⁰ Martin Ford, “*Architects of Intelligence. The truth about AI from the people building it*”, Packt Publishing, 2018, p.13.

⁷¹ *Ibidem*, pp.376/377.

⁷² Tal como a própria empresa refere no seu Website: “*In the coming years, AI – and ultimately artificial general intelligence (AGI) – has the potential to drive one of the greatest transformations in history*”, Google DeepMind, “*Our Vision*”, disponível na internet em <https://deepmind.google/about/>.

⁷³ *Ibidem*, p.171.

⁷⁴ Martin Ford, “*Rule of the Robots. How Artificial Intelligence Will Transform Everything*”, tradução por Luís Santos, “*O Futuro da Inteligência Artificial. Uma transformação que já está a acontecer*”, p.151.

2.3.3. Artificial Super Intelligence

O terceiro estágio é o mais avançado de todos, uma Super Inteligência Artificial. O que aqui se pretende é que as máquinas consigam não só igualar o nível de capacidade e inteligência humana, como as superem em todas as maneiras possíveis. A ASI visa superar a inteligência e habilidade do cérebro humano ⁷⁵.

Tal como referido anteriormente, ainda não atingimos sequer o nível da Inteligência Artificial Geral (AGI), mas Martin Ford refere que poderá ser através desta inovação que eventualmente se atingirá o nível da Superinteligência. Ao criar-se AGI, o sistema inteligente pode evoluir e tornar-se ainda mais inteligente do que o esperado, chegando mesmo ao ponto de superar qualquer capacidade intelectual do ser humano, ou até mesmo de ele próprio criar uma ASI ⁷⁶.

De acordo com Nick Bostrom, professor na Faculdade de Filosofia da Universidade de Oxford, Superinteligência pode ser definida como “*qualquer intelecto que exceda em muito o desempenho cognitivo dos seres humanos em praticamente todos os domínios de interesse.*” ⁷⁷. Contudo, o autor vai mais longe e, por ser uma definição um pouco vaga, apresenta três supercapacidades, três formas de ver a Superinteligência: a Velocidade, a Coletividade e a Qualidade.

A velocidade da ASI, é um sistema que consegue fazer tudo aquilo que o intelecto humano consegue, mas de uma maneira muito mais rápida (“*by multiple orders of magnitude*”) ⁷⁸. Para exemplificar esta supervelocidade, é dado o exemplo de uma emulação de um cérebro, em *hardware* próprio, 10.000 vezes mais rápido que o cérebro biológico, permitindo que o sistema consiga ler um livro de *n* páginas em meros segundos, ou então escrever uma tese de doutoramento numa tarde ⁷⁹.

No que diz respeito à Coletividade da ASI, o autor define-a como um sistema composto por um vasto número de pequenos intelectos de maneira que a *performance* geral do sistema, supere largamente a de qualquer sistema cognitivo atual, em vários domínios ⁸⁰. Desta maneira é possível obter um funcionamento mais eficaz e preciso na resolução de

⁷⁵ IBM, “*What is Artificial Intelligence?*”, disponível na internet em: <https://www.ibm.com/topics/artificial-intelligence>.

⁷⁶ Martin Ford, “*Architects of Intelligence. The truth about AI from the people building it*”, Packt Publishing, 2018, p.13.

⁷⁷ Nick Bostrom, “*Superintelligence: Paths, Dangers, Strategies*”, Oxford University Press, 2016, p.27.

⁷⁸ *Ibidem*, p.64.

⁷⁹ *Ibidem*.

⁸⁰ *Ibidem*, p.65.

problemas, pois este método permite que se divida uma questão em várias partes e que se aborde cada uma delas com o seu intelecto de especialização. Como na construção de um vaivém espacial, onde cada engenheiro está encarregue de uma parte específica daquele meio de transporte ⁸¹.

Por fim, temos a Qualidade da ASI, entendida como um sistema que é pelo menos tão rápido como a mente humana, mas tem uma qualidade de inteligência largamente superior ⁸². Aqui faz-se uma comparação entre a qualidade da inteligência demonstrada pelo ser humano face a alguns animais que apresentam pequenos índices de inteligência, logo inferior em qualidade. Neste caso, seria a ASI a apresentar uma inteligência de qualidade claramente superior quando comparada à humana, na realização de determinadas tarefas ⁸³.

2.3.4. *Machine Learning*

Machine Learning ou Aprendizagem Automática, segundo o dicionário de computação de Oxford, é um subcampo da Inteligência Artificial que se preocupa com a construção de programas que consigam aprender com as diversas experiências a que serão submetidos ⁸⁴, ou seja, é a disciplina que se foca na criação de algoritmos que vão aprender com as suas experiências e, a partir desses conhecimentos, evoluir, corrigir deficiências nas suas execuções e melhorar os seus desempenhos.

Importa clarificar que com a expressão *Machine Learning*, não se quer dizer que a máquina aprenda da mesma maneira que um ser humano aprende. Muito pelo contrário, aqui a “aprendizagem” é apenas uma metáfora face ao nosso processo de aprendizagem normal ⁸⁵.

Arlindo Oliveira recua até 1950, analisando algumas das ideias pioneiras que referenciam algoritmos adaptativos e com capacidade para aprender com dados ⁸⁶. No seu artigo “Computing Machinery and Intelligence”, Turing já proponha a produção de um programa que simulasse a mente de uma criança ao invés da mente de um adulto, pois aquela, não só seria de mais fácil construção (“*rather little mechanism*”), como também poderia servir como um caderno em branco a aguardar conhecimento (“*like a note-book (...)*”).

⁸¹ *Ibidem*, pp.65/66.

⁸² *Ibidem*, p.68.

⁸³ *Ibidem*, p.68/69.

⁸⁴ Dictionary of Computing, Fourth Edition, Oxford University Press, 1996, p.288.

⁸⁵ Harry Surden, “Artificial Intelligence and Law: An Overview”, in *Georgia State University Law Review*, Vol.3, issue 4, 2019, p.1311.

⁸⁶ Arlindo Oliveira, *Ob. cit.*, p.59.

lots of blank sheets”), ou seja, para o cientista computacional, seria mais benéfico criar um programa que simulasse um cérebro infantil “vazio” de conhecimento que, através de um processo de aprendizagem e educação, fosse adquirindo conhecimento e aumentando a sua base de dados até se tornar num “cérebro adulto”⁸⁷. Seria um “*sistema adaptativo, que usasse aprendizagem automática para adquirir a capacidade de raciocinar e resolver problemas complexos que associamos à inteligência humana*”⁸⁸.

Portanto, o algoritmo adaptativo de *Machine Learning*⁸⁹ vai detetar padrões, que lhe podem ter sido introduzidos previamente, de maneira a conseguir automatizar certos atos ou até mesmo fazer previsões com base em informação armazenada e recolhida^{90 91}. Um exemplo de um algoritmo de *Machine Learning* encontra-se no filtro nos nossos *emails*, que reconhece o *spam* ou ofertas comerciais, redirecionando-as para as respetivas pastas⁹². Este método está também presente nos motores de busca, que podem preencher automaticamente o pretendido pelo utilizador ou fazer recomendações relacionadas com o assunto em causa⁹³. Sistemas de reconhecimento visual começam também a ganhar mais força no campo do *Machine Learning*⁹⁴. As recomendações de produtos feitas pela Amazon, as músicas ou

⁸⁷ Alan Turing, *Ob. cit.*, p.456.

⁸⁸ Arlindo Oliveira, *Ibidem*.

⁸⁹ O algoritmo funciona como uma instrução passo-a-passo que vai ser codificada como *software* para que, desta forma, possa ser lida e analisada pelo computador. (Research Group on the Regulation of the Digital Economy, “*Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective*”, Max Planck Institute for Innovation and Competition, Research Paper no.19.13, 2019, p.4);

⁹⁰ Harry Surden, “Machine Learning and Law”, in *Washington Law Review*, Vol. 89, 2014, p.89.

⁹¹ Arlindo Oliveira exemplifica o conceito de aprendizagem automática da seguinte maneira: “*Suponhamos um sistema muito simples que recebe como entrada um único número e gera na sua saída um único número, que depende do primeiro. Se, a este sistema forem mostrados diversos exemplos da correspondência pretendida entre o número na entrada e o número na saída, o sistema poderá aprender, de forma aproximada, esta correspondência. Suponhamos que o número na entrada é a altura de uma pessoa e o número na saída é o seu peso. Se forem dados diversos exemplos ao sistema, de pares altura/peso, o sistema poderá aprender uma correspondência, aproximada, que dê como resultado o peso como função da altura.*”. Claro que este método não é infalível, e o autor faz menção à possibilidade de “erro”, mas é um método que, com mais dados, aumentará a sua percentagem de acerto, tornando-se cada vez mais preciso. (“(...) *dado um conjunto de dados na entrada, o objetivo é obter um programa que gere na saída uma estimativa do valor pretendido.*”). O método de *Machine Learning* é das armas mais fortes que a ciência tem ao seu dispor para poder atingir níveis mais elevados de automatização, tal como já é observável em cenários mais complexos como o dos veículos autónomos. “*Dotado dos algoritmos de aprendizagem automática adequados, um computador pode aprender a conduzir um veículo, simplesmente por análise do comportamento de um condutor humano (...)*”. Arlindo Oliveira, *Ob. cit.*, pp.60-63.

⁹² O algoritmo foi desenhado para detetar certas características específicas de emails de *spam*, e treinado para reconhecer esse padrão e abordá-lo de determinada maneira. (Harry Surden, *Ob. cit.*, pp.90-93).

⁹³ International Business Machines Corporation, “*Machine Learning*”, Design for AI, disponível na internet em <https://www.ibm.com/design/ai/basics/ml/>.

⁹⁴ “*Developments in ML has enabled us to supply pictures of, for example, a cat and over time, machines will begin to discern which pictures have cats in them from data it hasn't seen yet.*”, (*Ibidem*).

filmes que plataformas como a Netflix, Spotify e Youtube nos propõem ocasionalmente, são todas baseadas em *Machine Learning* ⁹⁵.

Da mesma maneira que o ser humano tem diversas formas de obter conhecimento (para se aprender um desporto não basta ler ou estudar sobre ele, é preciso praticá-lo regularmente), também as máquinas terão diferentes formas de “Aprendizagem”, como a Supervisionada, Não Supervisionada e a Por Reforço.

2.3.4.1. *Supervised Learning*

A aprendizagem supervisionada é um dos meios de aprendizagem automática mais usados na prática científica. O processo neste tipo de aprendizagem caracteriza-se pela introdução de vários tipos de informação rotulados de uma determinada maneira para que o sistema os armazene e utilize como parâmetros nas decisões futuras ⁹⁶. Portanto, vão ser disponibilizadas ao computador características específicas de algo, exemplos de imagens, resultados ou objetos, informação rotulada e devidamente identificada, para que, posteriormente, o sistema consiga fazer essa associação de forma autónoma.

Assim, um aspeto fundamental deste tipo de aprendizagem é a supervisão que será feita por um ser humano, ou seja, há intervenção humana principalmente na categorização da informação que será depois disponibilizada. “(...) *learning algorithms that learn to associate some input with some output, given a training set of examples of inputs x and outputs y . In many cases the outputs y may be difficult to collect automatically and must be provided by a human “supervisor” (...)*” ⁹⁷.

A *International Business Machines Corporation* exemplifica este tipo de aprendizagem comparando-o à de uma criança a aprender as formas geométricas. Num primeiro momento, é dada uma folha com as formas e respetivas designações e depois fornece-se algumas informações adicionais do tipo: se os círculos são redondos os triângulos

⁹⁵ *Machine Learning* é uma ferramenta muito poderosa e não é só dentro dos limites do *Online* que consegue ter presença. Os nossos *smartphones*, objetos que dificilmente são esquecidos por todos nós no dia a dia, funcionam com algoritmos de ML para quase todas as suas tarefas. “*The smartphone in your pocket (...) is learning about you. It uses Machine Learning to understand what you say, to correct your typos, to predict what you are going to do next and make suggestions. It learns your habitual routes from the GPS, and whether you tend to be late for meetings by comparing it with your calendar. It can even learn how you walk from its accelerometer. One of these days your smartphone is going to warn you and call 911 if it thinks you’re about to have a heart attack. So, Machine Learning can save your life.*”, Pedro Domingos, “*The Quest for the Master Algorithm*”, Tedx Talks Conference, 2016, disponível na internet em https://www.youtube.com/watch?v=qIZ5PXLVZfo&ab_channel=TEDxTalks.

⁹⁶ Research Group on the Regulation of the Digital Economy, *Ob. cit.*, p.7.

⁹⁷ Ian Goodfellow, Yoshua Bengio, Aaron Courville, “*Deep Learning*”, The MIT Press, Cambridge, 2016, p.136.

tem apenas três lados, podendo haver bastante variabilidade no seu tamanho. Informação adicional, ou mais específica, bem como alguns exemplos práticos são fundamentais para que não haja erros como, por exemplo, assumir que todos os triângulos têm os lados iguais, quando na verdade podem ser equiláteros, isósceles ou escalenos ⁹⁸. Dar a conhecer ao algoritmo o que é um cão, apresentando-lhe imagens onde o animal esteja presente (“*labeled “Dog”*”), mas incluir também várias imagens sem qualquer animal (“*labeled “No Dog”*”) ⁹⁹. O sistema, após assimilar e “aprender” toda esta informação rotulada, terá bases suficientes para a reconhecer sem muita dificuldade, variadas situações.

Faz-se ainda uma distinção com base no objetivo a atingir, sendo um “*classification problem*”, quando se pretende a previsão de uma coisa, um “*regression problem*”, quando se quer chegar a um valor numérico, e por fim um “*pattern recognition*”, quando se pretende reconhecer faces ¹⁰⁰.

Tal como referido anteriormente, este tipo de aprendizagem supervisionada é a mais comum, na prática, devido à “facilidade” de treino do sistema sendo, por isto, um método que se encontra principalmente nas grandes empresas comerciais como a Google, Amazon e Meta, que têm vastas quantidades de informação à sua disposição. Este método de aprendizagem é ainda observável em sistemas de tradução de linguagem, onde se faz a correspondência das palavras entre as línguas, e também, por exemplo, em sistemas de Radiologia que, recorrendo a imagens médicas, obtém diagnósticos através de informação previamente rotulada como “Cancro” ou “Não cancro” ¹⁰¹.

2.3.4.2. Unsupervised Learning

A Aprendizagem Não Supervisionada, contrariamente à anterior, é caracterizada pela falta de supervisão, ou seja, há uma redução substancial da intervenção humana. Aqui, não existe qualquer rótulo da informação introduzida no sistema, o que significa que o algoritmo é treinado para identificar características comuns ou semelhantes ou até mesmo identificar as diferenças da informação recolhida, sem qualquer auxílio (*labelling*) ¹⁰². Normalmente, este processo funciona através de agrupamentos (“*clusterings*”), onde o algoritmo vai dividir e agrupar a informação com base nas semelhanças ou diferenças concluídas ¹⁰³.

⁹⁸ International Business Machines Corporation, *Ibidem*.

⁹⁹ Martin Ford, “*Architects of Intelligence. The truth about AI from the people building it*”, p.11.

¹⁰⁰ International Business Machines Corporation, *Ibidem*.

¹⁰¹ Martin Ford, *Ibidem*, p.11/12.

¹⁰² Ian Goodfellow, *Ob. cit.*, pp.102 e 142.

¹⁰³ Stuart Russel, Peter Norvig, *Ob. cit.*, p.694/695.

Como não há necessidade de rótulos, a participação humana é diminuída, contudo, aumenta a importância da “*interpretação humana dos outputs*”¹⁰⁴, ou seja, a interpretação dos resultados.

Segundo o entendimento de Martin Ford este é dos primeiros tipos de aprendizagem que temos como seres humanos. Enquanto crianças, somos introduzidos aos primeiros aspetos da linguagem com base na informação que ouvimos dos nossos pais e familiares. Aprendemos diretamente da informação, não estruturada, recolhida do nosso meio ambiente¹⁰⁵.

Recorrendo ao exemplo da IBM, uma criança, sem qualquer indicação prévia, poderá agrupar as diversas formas desconhecidas baseando-se em critérios como a cor, o tamanho ou o formato¹⁰⁶. Ora, um poderoso algoritmo destes, permite-nos interpretar certos resultados ou formular algumas teorias como consequência desse “agrupamento”.

Este tipo de aprendizagem reforça a vertente autónoma do algoritmo pois está fortemente dependente das conclusões a que o *software* chega apenas com base na interação com o seu ambiente e observação dos dados que lhe são fornecidos.

É um importante método de aprendizagem que não só permite encontrar e escoar anomalias em quantidades vastas de informação, precisamente devido aos “*clusterings*” que realiza, como também vai permitir, com um adequado progresso, encurtar a distância para a obtenção de “*IA de nível humano*”^{107 108}.

2.3.4.3. Reinforcement Learning

Um algoritmo baseado numa aprendizagem por reforço é um algoritmo que vai aprender essencialmente através da tentativa e erro. Em nenhum momento é introduzida qualquer informação rotulada como correta ou incorreta, sendo o próprio *software* que chega à resposta certa após várias tentativas com determinados *feedbacks*¹⁰⁹. O algoritmo saberá que encontrou a solução pois será “recompensado” assim que a tiver obtido. Por outro lado, continuará à procura da resposta, se se deparar com um *feedback* negativo¹¹⁰.

¹⁰⁴ Research Group on the Regulation of the Digital Economy, *Ob. cit.*, p.8.

¹⁰⁵ Martin Ford, *Ibidem*, p.12.

¹⁰⁶ International Business Machines Corporation, *Ibidem*.

¹⁰⁷ Stuart Russel, Peter Norvig, *Ob. cit.*, p.695; e Martin Ford, *Ibidem*, p.12.

¹⁰⁸ É ainda apontada a possibilidade de existir uma *Semi-Supervised Learning* nos cenários onde as diferenças entre os dois tipos de aprendizagem não são tão claras. “*In semi-supervised learning we are given a few labeled examples and must make what we can of a large collection of unlabeled examples*”. Stuart Russel, Peter Norvig, *Ob. cit.*, p.695.

¹⁰⁹ Research Group on the Regulation of the Digital Economy, *Ob. cit.*, p.8.

¹¹⁰ Stuart Russel, Peter Norvig, *Ob. cit.*, p.830 e seguintes.

Portanto, a Aprendizagem por Reforço é sobre presentear as escolhas “certas” e punir os comportamentos “errados” ou não pretendidos.

É um método semelhante a treinar um cão para fazer determinados truques. Caso sente ou ladre quando for comandado para tal, será recompensado com um prémio ¹¹¹.

Estamos perante um método que exige muita informação e, principalmente, um longo treino com várias tentativas, para que se chegue ao melhor resultado possível. Para que isto aconteça é preciso um poderoso computador que consiga simular todas as rotas prováveis a uma rápida velocidade. Esta é uma das razões que torna este algoritmo de aprendizagem o ideal para o desenvolvimento de veículos autónomos, treinando-se o *software* num ambiente simulado e ensinando-lhe o que deve e não deve de fazer na estrada através de vários *feedbacks* ¹¹².

Um outro exemplo onde este algoritmo pode ter uma grande preponderância é no mundo dos videojogos, instruindo o sistema a prever, com várias jogadas de antecedência, qual a melhor estratégia a seguir para atingir a vitória ¹¹³.

A DeepMind da Google volta a ter aqui uma participação ativa com o desenvolvimento do AlphaGO, um sistema de Inteligência Artificial que combina redes neuronais profundas com algoritmos de pesquisa avançados para jogar o complexo jogo chinês de estratégia, Go (“*more complex than chess*”). Com recurso à aprendizagem por reforço, a empresa conseguiu criar um algoritmo que simulou milhares de jogos e milhões de jogadas possíveis “*each time learning from its mistakes*”, até que eventualmente conseguiu vencer um campeão mundial de Go ^{114 115}.

Ainda dentro dos esforços feitos pela DeepMind, de referir também a criação de um outro sistema, o AlphaZero, desta vez programado para aprender a jogar xadrez, shogi (xadrez japonês) e Go. AlphaZero começou por ser uma rede neuronal profunda juntamente com alguns algoritmos gerais que nada sabiam dos jogos para além das regras básicas, depois, através de um processo de tentativa/erro, o referido *Reinforcement Learning*, simulou milhões de jogos, aprendendo com as derrotas, empates e vitórias. Conseguiu assim

¹¹¹ Martin Ford, *Ibidem*, p.12.

¹¹² Martin Ford, *Ibidem*.

¹¹³ Stuart Russel, Peter Norvig, *Ob. cit.*, p.695.

¹¹⁴ Em outubro de 2015, AlphaGo realizou e venceu por 5-0 uma partida contra Fan Hui, três vezes campeão europeu do jogo. Menos de um ano depois, atinge outra vez a glória, ao vencer por 4-1 o campeão mundial Lee Sedol (vencedor de 18 títulos mundiais) num jogo lendário que fica recordado para a história como uma conquista que estava uma década à frente do seu tempo. Este evento deu origem a um documentário cheio de emoção, vencedor de vários prémios. (disponível na internet em https://www.youtube.com/watch?v=WXuK6gekUIY&t=4138s&ab_channel=GoogleDeepMind).

¹¹⁵ Google DeepMind, “AlphaGo”, disponível na internet em <https://deepmind.google/technologies/alphago/>.

superar o Stockfish ¹¹⁶ num jogo de xadrez, em apenas 4 horas, o Elmo ¹¹⁷ no Shogi, em apenas 2 horas, e por fim também já superou o seu precedente em Go, o Alphago, após 30 horas ¹¹⁸.

2.3.5. Deep Learning

A Aprendizagem Profunda ou “*Deep Learning*” é um subcampo do *Machine Learning*, o que significa que tudo o que foi exposto em relação ao conceito anterior, também é aqui aplicável. As principais diferenças encontram-se na *performance* dos algoritmos e no método utilizado por estes ¹¹⁹.

Um algoritmo de *Deep Learning* executa o processamento de informação através de uma rede neuronal artificial, uma *Artificial Neural Network* (ANN). Uma ANN é composta por um conjunto de neurónios artificiais em várias camadas que estão comprimidas entre outras duas: camada de entrada (“*input layer*”) e a camada de saída (“*output layer*”) ¹²⁰.

Por outras palavras, ANN pretende recriar aquilo que o cérebro biológico faz através dos neurónios e sinapses ¹²¹. Não é tarefa fácil pois o “*funcionamento de um neurónio (...) é complexo e multifacetado*” e o seu processo de comunicação biológico é feito através de sinapses, contudo, os diversos avanços observados ao longo dos anos, fizeram com que fosse possível que, com os métodos matemáticos e *hardware* adequado, se interligassem os

¹¹⁶ É considerado um dos mais poderosos motores de xadrez existentes, sendo ainda gratuito e de código aberto. Foi desenvolvido por Tord Romstad, Marco Costalba e Joona Kiiski, para ser utilizado como uma poderosa ferramenta à disposição dos jogadores, dada a sua alta capacidade de processamento e aprendizagem. (Stockfish, disponível na internet em <https://stockfishchess.org/about/>).

¹¹⁷ Um poderoso Software de Shogi desenvolvido por Makoto Takizawa.

¹¹⁸ Google DeepMind, “*AlphaZero*”, disponível na internet em <https://deepmind.google/discover/blog/alphazero-shedding-new-light-on-chess-shogi-and-go/>.

¹¹⁹ International Business Machines Corporation, “*Machine Learning*”, Design for AI, disponível na internet em <https://www.ibm.com/design/ai/basics/ml/>.

¹²⁰ International Business Machines Corporation, “*Neural Networks*”, disponível na internet em <https://www.ibm.com/topics/neural-networks>.

¹²¹ Paulius Cerka, Jurgita Grigienė, Gintarė Sirbiškytė, “Liability for damages caused by artificial intelligence”, in *Computer Law & Security Review*, n. °31, 2015, p.380.

neurónios artificiais com “links” numéricos ¹²² que transportam toda a informação pelas variadas camadas da ANN até chegarmos a determinados *outputs* ¹²³.

Esta é uma ferramenta poderosa que permite a classificação e agrupamento de informação a altas velocidades.

Importante referir que teremos diferentes tipos de *Artificial Neural Networks*: *Feedforward Neural Networks* (FNN); *Convolutional Neural Networks* (CNN) e *Recurrent Neural Networks* (RNN) ¹²⁴.

As Redes Neurais *Feedforward*, ou multi-layer perceptrons (MLPs), são as mais comuns e que se caracterizam pelas duas camadas de *input* e *output*, havendo entre elas camadas ocultas, processando a informação numa direção (*input*>*output*). Assim são basilares para o processamento de linguagem natural ¹²⁵.

Por sua vez as Redes Neurais Convolucionais são semelhantes às FNN, mas destacam-se pela sua *performance* superior com imagem, diálogo ou sinais de áudio ¹²⁶.

Por fim as Redes Neurais Recorrentes, são identificadas pela utilização de dados sequenciais ou de séries temporais, sendo por isso usualmente utilizadas para questões que persistem no tempo, como o reconhecimento de fala, a captura de imagens ou o processamento de linguagem natural ¹²⁷.

¹²² Estes links que interligam os neurónios artificiais estão “*modelados por um peso, um número, que é tanto maior quanto mais forte for a interligação*” (Arlindo Oliveira, *Ob. cit.*, pp.60-63). De uma maneira mais simplificada, cada neurónio artificial tem associado a si um peso e um limite, se o *output* de um neurónio individual for superior ao seu limite, então o neurónio ativa-se e manda informação para a próxima camada da rede, caso contrário não há ativação e conseqüentemente não haverá transmissão de *data*. (Stuart Russel, Peter Norvig, *Ob. cit.*, pp.728/729). Todas as informações introduzidas no sistema (*input*), têm pesos correspondentes com as respetivas características, sendo assim, com o treino adequado, o algoritmo de *deep learning*, após receber a informação, analisa e compara com a sua base de dados, multiplicando os pesos em causa e ativando assim os neurónios que o levam logicamente a um determinado resultado. No final resta fazer uma comparação das conclusões com os rótulos previamente introduzidos (no caso de uma *Supervised* ou *semi-supervised learning*), ou então validar/invalidar a conclusão com um *feedback* positivo/negativo (no caso da *Unsupervised learning*).

Nas palavras de Pedro Domingos, “(...) *in essence it just consists of repeatedly strengthening or weakening the connection between each pair of neurons in order to improve de accuracy of the predictions of the algorithm.*”, (Pedro Domingos, “*The Quest for the Master Algorithm*”, Tedx Talks Conference, 2016, disponível na internet em https://www.youtube.com/watch?v=qIZ5PXLVZfo&ab_channel=TEDxTalks; e David Rumelhart, Geoffrey Hinton, Ronald Williams, “*Learning Representations by back-propagating errors*”, in *Nature*, Vol.323, outubro, 1986, pp. 533-536).

¹²³ Stuart Russel, Peter Norvig, *Ob. cit.*, pp.728/729.

¹²⁴ International Business Machines Corporation, “*Neural Networks*”, disponível na internet em <https://www.ibm.com/topics/neural-networks>.

¹²⁵ *Ibidem*.

¹²⁶ International Business Machines Corporation, “*What are Convolutional Neural Networks?*”, disponível na internet em <https://www.ibm.com/topics/convolutional-neural-networks>.

¹²⁷ International Business Machines Corporation, “*What are Recurrent Neural Networks?*”, disponível na internet em <https://www.ibm.com/topics/recurrent-neural-networks>.

2.4. Inteligência Artificial nos dias de hoje

A área da Inteligência Artificial sofreu, ao longo dos anos, incontáveis mudanças e revoluções, atingindo marcos históricos e apresentando ao mundo novas tecnologias que antes pareciam pura ficção, contudo, mesmo com este já longo caminho repleto de sucessos e falhanços, ainda estamos longe de atingir todo o seu verdadeiro potencial.

É pertinente saber se, nos dias de hoje, o famoso Teste de Turing encontrou resposta para o seu dilema. Por outras palavras, será que podemos afirmar que a humanidade já alcançou o nível de Inteligência Artificial capaz de superar esse teste empírico? A resposta pode depender da exigência com que se realiza o Teste de Turing. O conceituado cientista fez, no seu artigo, uma previsão sobre o futuro das máquinas, afirmando que “*in about fifty years’ time it will be possible to programme computers (...) to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning.*”¹²⁸. Foi uma visão muito otimista, mas que hoje não está muito longe da realidade¹²⁹. Se colocarmos como interrogador um homem comum, sem qualquer conhecimento especializado nesta matéria, é bem possível que este tenha alguma dificuldade em distinguir a máquina do ser humano¹³⁰. Por outro lado, um juiz rigoroso (“*sophisticated judge*”) e conhecedor dos conceitos associados à temática, já será mais difícil de enganar¹³¹.

Portanto, a resposta à pergunta anterior parece ser subjetiva, dependendo de como queremos interpretar o teste. Numa vertente mais literal, a verdade é que já temos vários sistemas que simulam e mantêm conversas com o ser humano, tornando-se assim cada vez mais indistinguíveis (mesmo que não em termos absolutos, pelo menos aos olhos de interrogadores especializados)¹³². Numa vertente mais simbólica, a criação de um sistema

¹²⁸ “*dentro de cerca de 50 anos, será possível programar computadores (...) para fazê-los jogar tão bem o jogo da imitação, que um interrogador médio não terá mais de 70% de probabilidade acertar a identificação, após cinco minutos de interrogatório*”, Alan Turing, *Ob. cit.*, p.442.

¹²⁹ Arlindo Oliveira, *Ob. cit.*, p.44.

¹³⁰ Recorrendo a um exemplo prático já referenciado anteriormente, *ChatGPT* ou *Gemini*, que como exemplos de IAs conversacionais podem responder a perguntas e até mesmo manter uma espécie de conversa contínua com alguém, desde que tenham as informações dentro das suas bases de dados. Contudo, estes sistemas não só se encontram limitados pelas suas bases de dados (algumas ainda desatualizadas, não tendo informações do ano atual), como também estão programados para não induzir o utilizador em erro, principalmente quando se apresentam como *chatbots* por não conseguirem satisfazer o pedido.

¹³¹ Stuart Russel, Peter Norvig, *Ob. cit.*, p.1021.

¹³² Joseph Weizenbaum criou em 1960 o sistema ELIZA, que usando frases pré-escritas, respondia às questões que lhe eram feitas. Mesmo sem qualquer compreensão, o sistema conseguiu enganar muitos utilizadores. (“*(...) ELIZA foi um dos primeiros sistemas a passar o um teste de Turing, embora um teste administrado em condições muito específicas e pouco exigentes.*” Arlindo Oliveira, *Ob. cit.*, p.54).

E muitos outros sistemas que conseguem, em certa medida, superar o teste de Turing (convencendo uma pequena parte dos juízes), podem ser encontrados entre a lista de vencedores do “Loebner Prize Competition”,

que estabeleça um diálogo com uma pessoa e apresente uma compreensão e naturalidade indistinguíveis de qualquer outro ser humano já parece ser um cenário mais distante.

A criação de sistemas com uma grande capacidade de processamento de dados tornou-se comum nos dias de hoje. Parece que há sempre um sistema ou algoritmo que consegue fazer tudo, ou quase tudo, de forma mais célere e automática. Esta cultura de IAs é observável por exemplo na existência do CAPTCHA (*Completely Automated Public Turing Teste to tell Computers and Humans Apart*), que como o nome indica é um teste de Turing público para distinguir computadores de pessoas, habitualmente utilizado em *websites* comerciais e que serve como medida de segurança no combate ao crescente número de *bots* na internet, que são uma clara ameaça de *spam*, ou pior, de fuga de informação.

Contudo, no final do dia, enquanto humanidade ainda não atingimos, artificialmente, o nível da inteligência humana, ainda não atingimos a inteligência a que Turing se referia no seu artigo “*Can Machines Think?*”.

Ainda na esfera das previsões feitas por alguns cientistas, também Marvin Minsky e Herbert Simon, juntamente com outros investigadores, quiseram deixar a sua opinião, e em 1960, afirmaram que dentro de três décadas, a humanidade conseguiria desenvolver sistemas de inteligência comparáveis à humana que consigam realizar qualquer função desempenhada pelo homem ¹³³. Mais uma vez, esta previsão foi demasiado otimista, pois elegeu os anos 90 como uma era mais tecnológica do que aquilo que efetivamente foi. Por outro lado, em pleno século XXI, ainda não temos um sistema que recrie, de forma absoluta, a inteligência humana, mas já temos algoritmos que processam informação e realizam tarefas de forma mais eficaz e célere que o próprio ser humano.

Podemos dizer que do ponto de vista prático, a ciência já conseguiu igualar e superar a *performance* humana, mas no que respeita à recriação artificial de todas as capacidades do cérebro humano, é ainda uma tarefa de difícil realização.

2.5. Futuro da Inteligência Artificial

Falar do futuro neste tipo de áreas científicas como a da Inteligência Artificial é sempre uma tarefa de difícil execução, em grande parte devido à enorme complexidade do tema em questão, mas também por causa da imprevisibilidade das exigências e necessidades

uma competição que premiava programas de computador conforme o seu grau de semelhança com o ser humano.

¹³³ Arlindo Oliveira, *Ob. cit.*, p.54.

do ser humano. Ainda assim, muitos são corajosos ao ponto de tentar dar a sua opinião quanto a muitos destes dilemas.

Na compilação de conversas realizada por Martin Ford, um dos capítulos finais tem o título de “*When will human-level AI be achieved? Survey Results*”, e é onde o autor faz um resumo das inúmeras previsões, dadas por alguns investigadores, sobre quando acham que a humanidade atingirá o nível da Inteligência Artificial Geral (AGI), “*human-level AI*”. Os palpites são diversos, variando entre os mais otimistas, que apontam para anos mais próximos, como 2029 ou 2036, e os mais pessimistas, ou talvez realistas, que já colocam este tipo de tecnologia a surgir apenas daqui a uma centena ou mais de anos, como 2188 ou 2200, por exemplo ¹³⁴.

A criação de AGI ou o tornar as máquinas mais “humanas” não é o único foco da ciência, muito pelo contrário, é uma das várias ramificações que pertence ao raio de investigação da área da Inteligência Artificial. Fora isto, muitas outras inovações estão a ser projetadas e desenvolvidas.

A *Neuralink Corporation*, uma empresa cofundada pelo bilionário Elon Musk ¹³⁵, está neste momento a desenvolver *Brain-Computer Interfaces* (BCI), ou por outras palavras, implantes cerebrais que permitem controlar dispositivos eletrónicos (computadores ou telemóveis) através da interpretação da atividade neuronal de uma pessoa ¹³⁶. Após o tremendo sucesso dos BCIs em testes realizados em animais, a empresa obteve autorização para iniciar os testes com seres humanos ¹³⁷, dando assim início ao “PRIME Study” (*Precise Robotically IMplanted Brain-Computer interfAcE*), que pretende obter significativos avanços médicos na luta contra lesões da espinhal medula e contra a Esclerose Lateral

¹³⁴ Martin Ford, *Ob. cit.*, pp.528 e 529.

¹³⁵ Juntamente com um grupo de cientistas e engenheiros como o neurocirurgião Dongjin Seo, Max Hodak um cientista de dados e Paul Merolla, engenheiro eletrónico, entre outros.

¹³⁶ Neuralink Corporation, disponível na internet em <https://neuralink.com/>.

¹³⁷ A 29 de janeiro de 2024, a empresa anuncia a realização, com sucesso, do procedimento de implementação no primeiro ser humano (Noland Arbaugh, 29 anos, paralisado num acidente de mergulho). Apesar da escassa informação, Elon Musk demonstrou-se bastante ansioso e otimista face à recuperação do paciente e a potenciais resultados desejados, enaltecendo que, após algum treino, este conseguia controlar um rato de computador com apenas o pensamento. É com certeza um assunto precoce e com algum desconhecimento público, mas que se encontrará no foco de muitos, nos anos que se avizinham. Reuters, “*Future of Health*”, <https://www.reuters.com/business/healthcare-pharmaceuticals/neuralinks-first-human-patient-able-control-mouse-through-thinking-musk-says-2024-02-20/>. Tal como é de se esperar em qualquer inovação tecnológica, já foram confirmados alguns desafios inesperados, contudo, o procedimento parece estar a levar o seu curso natural, permitindo ao seu utilizador uma utilização quase que telepática do seu computador (CNN, <https://cnnportugal.iol.pt/noland-arbaugh/neuralink/primeiro-paciente-com-chip-cerebral-da-neuralink-filmado-a-jogar-xadrez-online/20240321/65f6e5f6d34e8d13c9b91041;>). É previsível que nem tudo se desenrole sem mais problemas, mas não restam dúvidas quanto ao potencial uso futuro desta poderosa tecnologia.

Amiotrófica (ALS) ¹³⁸. Num primeiro momento, o recrutamento para a fase de testes visa apenas pessoas que não tenham qualquer movimento, ou movimentos reduzidos, mas parece correto afirmar que, caso se atinja o sucesso com esta nova experiência, num futuro próximo este tipo de tecnologia seja disponibilizado ao público como o mais recente e topo de gama “complemento de *smartphone*”. Esta pesquisa não só poderá ajudar imensas pessoas com problemas de saúde complicados, melhorando-lhes a qualidade de vida, como poderá permitir um melhor conhecimento das respetivas condições médicas bem como permitir um melhor entendimento do cérebro humano, benéfico para o desenvolvimento das restantes áreas da IA e da medicina. ¹³⁹. Este é, sem dúvida, um dos projetos atuais mais ambiciosos de sempre.

Muitos outros são os projetos em andamento para os próximos pares de anos, e, ainda mais, são os constantes debates sobre o alcance e poder destas novas tecnologias. É verdade que uma sociedade mais tecnológica e automatizada tem os seus benefícios para a humanidade, mas também não deixa de ser verdade que um futuro desses não se traduz só em vantagens.

Um grande nível de automatização artificial permite uma gestão mais célere, cómoda e eficaz da vida pessoal e profissional, transformando grandes sequências de atos em meros cliques de alguns botões. O que antes implicava longos períodos de dedicação a investigação ou preparação de estudos, dados, organização de informação ou realização de cálculos, pode agora ser delegado, quase que de forma absoluta, a um sistema com um algoritmo que foi criado, exclusivamente, para a realização dessa tarefa. Esta tremenda simplificação da vida humana é vista por muitos como uma das grandes vantagens da criação de sistemas de IA, que vão progressivamente acabando com um papel mais ativo do ser humano, convertendo-o num mero vigilante, um mero observador encarregue apenas da manutenção dos sistemas que fazem o trabalho todo. A criação de uma “*Post-Work Utopia*” ¹⁴⁰.

¹³⁸ Neuralink Corporation, “*Patient Registry*”, disponível na internet em <https://neuralink.com/patient-registry/>.

¹³⁹ Tal como a própria empresa refere na sua explicação “*This research will be the first of its kind to be performed in people and may help us find safer, more effective ways to implant and use our BCI to potentially restore and enhance computer control and other capabilities.*” Neuralink Corporation, “*Patient Registry*”, disponível na internet em <https://neuralink.com/patient-registry/>.

¹⁴⁰ “*The use of machines and computer programs has obviated the need for human work. Members of that society are liberated from a life mostly filled with hard industrial labor or mindless office jobs, and people have arrived at a productive, happy, and more meaningful level of existence. They are able to spend their time engaging in arts, music, literature, science, and in rich social relationships. Robots have taken over agricultural and industrial production of all commodities people need and desire. All goods are produced without the help of humans.*”. Steven John Thompson, “*Machine Law, Ethics, and Morality in the Age of Artificial Intelligence*”, Engineering Science Reference (IGI Global), 2021, p.207.

Em contraste com esta ideia de uma vida mais simples persistem as preocupações éticas e morais relativamente a uma utilização excessiva destes sistemas de IA, por exemplo, no meio de trabalho. Eventualmente, a grande maioria dos empregos realizados por seres humanos passarão a ser exercidos, exclusivamente, por *softwares* inteligentes levando a uma irradiação em massa das oportunidades de trabalho de grande parte da população. Este desemprego tecnológico (“*technological unemployment*”) pode resultar em vários problemas de desigualdades sociais, visto que o lucro ficará retido nos grandes empresários que não terão mais necessidade ou gastos com a mão de obra humana ¹⁴¹.

Uma outra questão que também tem bastante relevância nestes debates é a relativa à segurança e à privacidade. Como sociedade todos queremos viver em liberdade, mas com garantias que assegurem a nossa segurança. Tem-se verificado um aumento considerável na investigação e criação de sistemas de reconhecimento facial, dotados de algoritmos que, a partir de informação recolhida ou armazenada, identificam determinadas pessoas ou suspeitos, estudam comportamentos para prever e prevenir práticas de atos criminosos, entre outras funções. Esta pode ser uma ferramenta verdadeiramente útil para as forças de segurança, mas terá sempre de ter uma rigorosa regulação para que não se banalize um importante direito como a privacidade, sem nunca esquecer o potencial perigo de viés raciais e discriminatórios dos algoritmos ^{142 143}.

O futuro é claramente incerto, mas antes do próximo grande marco histórico no mundo da Inteligência Artificial, muitas são as questões que têm de ser debatidas, não só do ponto de vista legal, no caso da responsabilidade civil destes sistemas e dos Direitos Humanos com os quais possam conflitar, mas também do ponto de vista ético e social. É importante estabelecer-se boas bases e regras, para que o desenvolvimento ocorra, mas de uma maneira que nos traga mais benefícios do que desvantagens.

¹⁴¹ Steven John Thompson, *Ibidem*, pp.205/206.

¹⁴² A China é um exemplo que demonstra o perigo de uma sociedade excessivamente controladora, com uma rede enorme de vigilância impulsionada por sistemas de IA de reconhecimento facial e outras tecnologias mais avançadas. Uma sociedade super tecnológica, na frente mundial das descobertas em matérias de Inteligência Artificial, que funciona na sua base com um sistema de classificação social geral para os seus cidadãos, concedendo vantagens para quem apresenta classificações positivas, ou penalizando quem se encontra fora da “norma” tipicamente social. Martin Ford, “*Rule of the Robots. How Artificial Intelligence Will Transform Everything*”, p.234. Esta realidade alavancada por um governo controlador e fortemente alicerçado em sistemas inteligentes não deve nunca ser um objetivo, mas sim um aviso para se fazer melhor.

¹⁴³ Valentina Golunova, “Artificial Intelligence and the Right to Liberty and Security”, in *Artificial Intelligence and, Human Rights*, edited by Alberto Quintavalla e Jeroen Temperman, Oxford University Press, 2021, p.45

2.6. Inteligência Artificial e o Direito

O impacto que a Inteligência Artificial está a ter em diversos setores da sociedade é tremendo e o Direito não foge a essa regra.

Com o contínuo e rápido crescimento tecnológico temos dois grandes problemas que se podem colocar na área jurídica. O primeiro está relacionado com a modernização do setor em si, o segundo diz respeito aos novos desafios legislativos provocados por uma sociedade cada vez mais tecnológica e automatizada.

O surgimento das novas tecnologias da informação e comunicação impulsionaram uma modernização natural do direito como um todo. A criação dos computadores e posteriormente do correio eletrónico, bem como de outras formas de prestar informações e realizar comunicações, culminou na natureza eletrónica dos processos tal como os conhecemos hoje, deixando para trás a regra do processo em suporte físico ¹⁴⁴.

A tramitação eletrónica dos processos, a criação de sistemas informáticos de suporte à atividade dos tribunais como o CITIUS e o SITAF, são fortes marcas da crescente desmaterialização do direito que trazem imensos benefícios, entre os quais a criação de uma justiça mais transparente e acessível, permitindo com grande facilidade consultar e conhecer-se os processos, sentenças e decisões; também comporta benefícios ambientais por se estar a reduzir substancialmente o uso de papel nos tribunais; de igual modo, apresenta uma melhoria no armazenamento e organização de dados e processos, visto que guardar centenas de milhares de informações em formato digital não requer o mesmo espaço físico que o arquivamento material; e por fim pode verificar-se um aumento da celeridade, agilização dos processos até certo nível, pois a tramitação é mais simples e, com um circuito eletrónico e automático, há uma dispensa de atos burocráticos intermédios alheios à decisão da causa ^{145 146}.

¹⁴⁴ De acordo com o artigo 132.º do CPC, o processo tem natureza eletrónica e a sua tramitação, bem como a prática de atos escritos, ocorrerá através do sistema informático adequado. Tramitação dos processos regulamentada pela Portaria n.º 280/2013, de 16 de agosto para os processos judiciais e pela Portaria n.º 380/2017, de 19 de dezembro para a jurisdição administrativa e fiscal.

¹⁴⁵ eTribunal CITIUS disponível na internet em <https://www.citius.mj.pt/portal/Faq.aspx>.

¹⁴⁶ Como resultados da desmaterialização do direito temos vários exemplos de sucesso ao nível nacional seja dos já mencionados sistemas informáticos de suporte à atividade dos tribunais judiciais e administrativos, CITIUS e SITAF, como de muitos outros como por exemplo, o e-Leilões, uma plataforma desenvolvida pela Ordem dos Solicitadores e dos Agente de Execução para a realização da venda de bens através de leilão eletrónico (e-leilões disponível na internet em <https://www.e-leiloes.pt/faqs.aspx>).

Ao nível europeu também existe um grande movimento de desmaterialização do direito podendo encontrar-se no Portal Europeu da Justiça menção a uma listagem vasta de projetos já desenvolvidos e outros ainda em estado de desenvolvimento, pretendendo não só automatizar mais o direito como também reforçar o relacionamento tecnologia-direito. (Portal Europeu da Justiça, disponível na internet em <https://e-justice.europa.eu/home?plang=pt&action=home>). Um dos exemplos de maior sucesso é o e-CODEX como uma ferramenta à disposição dos Estados-Membros que lhes permite estabelecer uma rede de comunicações

Para além da grande desmaterialização do processo, a implementação de novas tecnologias, principalmente da Inteligência Artificial como um instrumento de auxílio às práticas jurídicas, também tem sido tema de muito debate. Pode estar em causa um auxílio mais simples ao nível da análise de factos e informação e até mesmo da interpretação jurídica, ou um auxílio mais avançado no que diz respeito à tomada de decisões, ou seja, a eventual criação de juízes robôs ¹⁴⁷.

Por um lado, apesar de incerto, não parece que a curto e até mesmo médio/longo prazo, se observe um movimento de substituição do juiz por um *software* sem qualquer intervenção humana, (“*the concept of a Robo-Judge, making decisions in most areas of law without human intervention is not feasible*” ¹⁴⁸), por outro lado, a criação de algoritmos ou sistemas que visam auxiliar e complementar o processo de decisão do juiz começam a ganhar cada vez mais força ^{149 150}.

Por sua vez, cresce a necessidade de adaptação do direito aos variados desafios que vão surgindo, uma adaptação a um mundo cada vez mais tecnológico. Tal como este trabalho almeja, é fundamental saber quais serão as mutações que o direito civil terá de sofrer para responder de maneira idónea às questões que se colocam no âmbito da responsabilidade, seja com veículos autónomos, com sistemas de Inteligência Artificial ou outro tipo de nova tecnologia.

E não ficamos por aqui, também do ponto de vista do direito laboral, penal, administrativo e, eventualmente, constitucional, terá de haver correções às regulações tradicionais para que, deste modo, o direito se mantenha atual e pronto a responder às necessidades de uma sociedade diferente daquela que hoje conhecemos.

interoperável, segura e descentralizada entre os respetivos sistemas eletrónicos nacionais, tanto em questões de matéria cível como penal (e-Codex, disponível na internet em <https://www.e-codex.eu/projects>).

¹⁴⁷ Elisa Alfaia Sampaio e Paulo Jorge Gomes, *Ob., cit.*, p.209 e seguintes.

¹⁴⁸ John Zeleznikow, “Using Artificial Intelligence to support legal decision-making – Is the software a friend or a foe?”, em *Inteligência Artificial e Robótica: Desafios para o direito no século XXI*, GESTLEGAL, 1ª edição, p.21-43, 2022, pp.23/24.

¹⁴⁹ Já é bastante comum ver a utilização de sistemas de IA ou assistentes jurídicos virtuais que prestam auxílio jurídico ao fornecer uma procura mais eficiente, rápida e automática de legislação ou jurisprudência com pouco ou nenhum esforço. Alguns exemplos são o “Alpaca Law”, sistema este composto por um algoritmo de pesquisa que permite encontrar vários precedentes de uma base de dados de jurisprudência atualizada diariamente. Mais, este assistente disponibiliza também uma função de interpretação e análise de acórdãos, podendo simplificá-los em meros segundos. (Alpaca Law, disponível na internet em <https://alpacaLaw.com/sobre>).

Outro exemplo semelhante encontra-se no Apura, um software que permite a procura inteligente de jurisprudência nacional (Apura, disponível na internet em <https://apura.ai/>);

Apesar de pioneiros e com bastante utilidade, são ainda trabalhos focados na maior agilização e automatização da recolha e análise de dados relevantes para o caso, afastando-se assim do cenário mais complexo da decisão por sistemas artificiais.

¹⁵⁰ Elisa Alfaia Sampaio e Paulo Jorge Gomes, *Ob., cit.*, p.209.

Assim, parece correto afirmar que o direito não só está em constante mudança por implementar diferentes tecnologias que modernizam o seu *modus operandi*, como também está, cada vez mais, a cooperar com sistemas inteligentes e algoritmos que auxiliam tanto advogados como juízes na prática das suas atividades jurídicas.

Por fim, lança-se a grande questão: Que resposta vai dar o direito quando se deparar com os danos causados pelos sistemas de inteligência artificial? Quem responderá pelos danos causados? O criador do algoritmo ou do *software*? A empresa que utiliza o sistema de forma comercial? O fabricante da Inteligência Artificial? O próprio sistema de IA? ¹⁵¹. A que título se verificará essa responsabilidade? Será sobre isto que nos debruçaremos.

¹⁵¹ José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, in *Revista de Direito Comercial*, edição de 2020, pp.78 e 79.

3. Responsabilidade Civil por danos causados por Sistemas de Inteligência Artificial

No presente capítulo faremos a exposição do problema central de que se ocupa esta dissertação: Quem vai responder pelos danos causados pelos sistemas de Inteligência Artificial? Será que os regimes tradicionais de responsabilidade estão preparados para satisfazer as necessidades de uma realidade distinta da, até então, verificada?

Logicamente, este raciocínio vai implicar, num primeiro momento, um breve desenvolvimento sobre os regimes clássicos de responsabilidade, de maneira que, num momento posterior, se possa não só apresentar as dificuldades sentidas pelas particularidades dos sistemas de inteligência artificial, como também apreciar criticamente as alternativas que a doutrina apurou para lidar com estas novas circunstâncias.

3.1. Responsabilidade Civil – Considerações Gerais

A responsabilidade civil é uma fonte de obrigações baseada no princípio do ressarcimento dos danos, ou seja, é um “conjunto de factos que dão origem à obrigação de indemnizar os danos sofridos por outrem”¹⁵². Nas palavras de Almeida Costa, esta figura jurídica ocorre quando “uma pessoa deve reparar um dano sofrido por outra”¹⁵³.

Há uma divisão da responsabilidade civil em responsabilidade civil contratual (ou obrigacional) e em responsabilidade civil extracontratual (extra-obrigacional ou delitual). A primeira resulta da “falta de cumprimento das obrigações emergentes dos contratos, de negócios unilaterais ou da lei”¹⁵⁴. Estamos assim perante uma violação de um direito de crédito ou de uma obrigação em sentido técnico^{155 156}. A responsabilidade contratual está

¹⁵² Luís Menezes Leitão, “Direito das Obrigações. Vol. I – Introdução da Constituição das Obrigações”, Almedina, Coimbra, 2022, Reimpressão 2023, p.277.

¹⁵³ Mário Júlio de Almeida Costa, “Direito das Obrigações”, Almedina, Coimbra, 12ª edição revista e atualizada, 2010, reimpressão 2023, pp.517 e 518.

¹⁵⁴ João Antunes Varela, “Das Obrigações em Geral Vol. I”, Almedina, Coimbra, 2017, reimpressão 2022, p.519.

¹⁵⁵ Mário Júlio de Almeida Costa, *Ob. cit.*, p.539.

¹⁵⁶ Discute-se que, face à denominação da responsabilidade contratual, a expressão “obrigacional” seria mais completa, precisamente por estar em causa não apenas a violação de obrigações provenientes de vínculos contratuais, mas também de negócios unilaterais e da própria lei. Contudo, também se critica esta designação pela sua equivocidade, pois esta não distingue o “dever de prestar, tendente ao cumprimento da obrigação”, do “dever de indemnizar, correspondente ao seu não cumprimento”. Outra possibilidade surgiu como responsabilidade “negocial”, mas para além de restrita também tem “falta de tradição que a consagre”, ao contrário da expressão “contratual”. (Mário Júlio de Almeida Costa, *Ob. cit.*, p.539, nota de rodapé; e João Antunes Varela, *Ob. cit.*, p.520, nota de rodapé).

consagrada, no nosso Código Civil nos artigos 798.º e seguintes. Por sua vez, a responsabilidade extracontratual, prevista nos artigos 483.º e seguintes, deriva da “*violação de direitos absolutos ou da prática de certos atos, que embora lícitos, causam prejuízo a outrem.*”^{157 158}. Assim, a grande diferença entre estes dois tipos de responsabilidade reside na existência, ou não, de uma relação intersubjetiva prévia entre lesante e lesado. Quando estejam em causa a violação de direitos absolutos completamente desligados de qualquer vínculo jurídico prévio, estaremos perante um caso de responsabilidade delitual, por outro lado, será responsabilidade obrigacional quando se violem deveres emergentes de uma relação específica previamente existente entre as partes¹⁵⁹.

Podemos ainda apontar uma terceira via de responsabilidade denominada de responsabilidade pré-contratual que abrange as situações onde ocorrem as violações dos deveres de boa-fé¹⁶⁰. Na base desta modalidade reside a ideia de proteger os interesses e a confiança que as partes vão ganhando antes mesmo da celebração do contrato. De acordo com o n.º 1 do artigo 227.º, “*Quem negocia com outrem para conclusão de um contrato deve, tanto nos preliminares como na formação dele, proceder segundo as regras de boa-fé, sob pena de responder pelos danos que culposamente causar à outra parte.*” Nos dias de hoje, com um processo de formação dos contratos cada vez mais longo e complexo, seria inaceitável que não fosse juridicamente protegida a confiança e expectativa que as partes adquirem em relação à conclusão do contrato¹⁶¹.

Olhando novamente para o artigo 483.º do código civil constatamos o reconhecimento expresso de duas formas de responsabilidade civil extracontratual: a subjetiva e a objetiva. De acordo com o n.º 1 deste preceito legal, fica claro que o princípio geral é o da responsabilidade subjetiva, ou seja, existe obrigação de indemnizar daquele que agir com dolo ou mera culpa. De outro modo, o n.º 2 atribui caráter excecional à responsabilidade objetiva, só podendo existir obrigação de indemnizar, independentemente de culpa, nos casos legalmente previstos¹⁶².

¹⁵⁷ João Antunes Varela, *Ob. cit.*, p.520.

¹⁵⁸ Aqui também surgiu a temática das designações, podendo esta responsabilidade ser denominada de extra-obrigacional ou extranegocial. Assim sofre das mesmas críticas apontadas anteriormente às expressões para a responsabilidade contratual (Mário Júlio de Almeida Costa, *Ob. cit.*, p.539, nota de rodapé;). Surge também como responsabilidade delitual, que suscita algumas reservas pela imprecisão da noção de delito, visto que “*a responsabilidade civil se estende para além dos casos de delitos tipificados na lei penal*” (João Antunes Varela, *Ob. cit.*, p.520, nota de rodapé).

¹⁵⁹ Luís Menezes Leitão, *Ob. cit.*, p.280.

¹⁶⁰ *Ibidem*.

¹⁶¹ *Ibidem*, p.354.

¹⁶² João Antunes Varela, *Ob. cit.*, p.523.

Assim, na responsabilidade civil extracontratual ou delitual, temos por regra uma responsabilidade civil por factos ilícitos (subjativa), podendo existir em determinadas situações excepcionais uma responsabilidade pelo risco (objetiva). Há ainda espaço para uma responsabilidade civil por factos lícitos causadores de danos. Olhemos com mais detalhe os dois primeiros tipos, nos próximos subcapítulos.

3.2. Responsabilidade Civil Extracontratual

3.2.1. Responsabilidade Civil Extracontratual por factos ilícitos

O artigo 483.º do CC, não só determina como regra geral a responsabilidade civil subjativa, como também estabelece os pressupostos genéricos da responsabilidade delitual, pressupostos estes que são de verificação cumulativa. Dita o n.º 1 do artigo que “*Aquele que, com dolo ou mera culpa, violar ilicitamente o direito de outrem ou qualquer disposição legal destinada a proteger interesses alheios fica obrigado a indemnizar o lesado pelos danos resultantes da violação.*”. É possível então, apurar 5 pressupostos: o facto, a ilicitude, a culpa, o dano e por fim o nexo de causalidade entre o facto e o dano. Aprofundemos mais cada um dos requisitos.

Em primeiro lugar, é preciso um facto voluntário por parte do agente, ou seja, é necessário que o agente tenha um comportamento dominável ou controlável pela sua vontade ¹⁶³. Por um lado, este facto pode ser uma ação, um facto positivo, estando em causa um comportamento em que o agente por exemplo lese alguém com uma ofensa corporal ou se aproprie ou destrua património alheio. Por outro lado, também pode ser uma omissão, um facto negativo, onde o agente se abstém de agir. Deste modo, para que das omissões resulte a obrigação de indemnizar é preciso que haja “*por força da lei ou de negócio jurídico, o dever de praticar o ato omitido*” ¹⁶⁴. Portanto, mesmo que um ato omissivo não gere física ou materialmente o dano, é considerado a sua causa pois, a sua realização, quando há um dever

¹⁶³ Não são considerados factos voluntários os comportamentos que não digam respeito ao agente, ou seja, causas de força maior ou outras circunstâncias fortuitas que originem das forças naturais invencíveis (por exemplo tempestades, tremores de terra, ciclones, alguém ser impelido por explosões ou descargas elétricas). Ainda assim, mesmo os que sejam respeitantes ao agente podem não cumprir este requisito se ele estiver por alguma razão desprovido da capacidade de os dominar (se alguém produzir um dano por consequência de uma síncope cardíaca ou sendo vítima de coação). Luís Menezes Leitão, *Ob. cit.*, p.281 e João Antunes Varela, *Ob. cit.*, p.529.

¹⁶⁴ Cfr. artigo 486.º do CC.

jurídico de agir criado por um negócio jurídico ¹⁶⁵ ou por força da lei ¹⁶⁶, muito provavelmente ou seguramente “*teria impedido a consumação desse dano*” ¹⁶⁷.

De seguida, o pressuposto que se segue é o da ilicitude, ou seja, para além de um facto voluntário é necessário que se verifique uma violação da lei, pois há certos atos que mesmo sendo prejudiciais não geram a obrigação de indemnizar por não ser cometida qualquer infração legal (como por exemplo, a criação de uma indústria que verse sobre o mesmo produto de uma outra já existente localmente, mesmo que os interesses do dono de uma das indústrias sejam lesados, isso não gera a obrigação de indemnização por não se verificar qualquer violação da lei ¹⁶⁸).

De acordo com o artigo 483.º/n.º 1, temos uma cláusula geral limitada inspirada no direito alemão ¹⁶⁹, que faz uma determinação algo abrangente de quais os bens jurídicos cuja lesão desencadeia responsabilidade. Assim teremos ilicitude do facto praticado pelo agente quando este “*violar ilicitamente o direito de outrem*” ou, quando violar “*qualquer disposição legal destinada a proteger interesses alheios*” ¹⁷⁰. Na primeira modalidade de ilicitude encontramos a violação de direitos subjetivos, tipicamente ofensas de direitos absolutos, como por exemplo os bens jurídicos pessoais (a vida, o corpo, a saúde ou a liberdade ¹⁷¹), os direitos reais, os direitos de propriedade industrial, os direitos de autor, os direitos de personalidade ¹⁷², os direitos familiares de natureza patrimonial (não se aplicando aos direitos familiares pessoais ¹⁷³). É relevante referir ainda que, ao se exigir a violação de um direito subjetivo em concreto, também se está a limitar a indemnização à “*frustração das*

¹⁶⁵ Por exemplo o dever que um membro do staff de um centro social de dia, escola ou hospital tem de vigiar qualquer idoso, criança ou utente. Luís Menezes Leitão, *Ob. cit.*, p.282.

¹⁶⁶ Como exemplo de algumas disposições onde a lei impõe determinados comportamentos temos os artigos 491.º Responsabilidade das pessoas obrigadas à vigilância de outrem, 492.º Danos causados por edifícios ou outras obras e 493.º Danos causados por coisas, animais ou atividades. Luís Menezes Leitão, *Ob. cit.*, p.282.

¹⁶⁷ João Antunes Varela, *Ob. cit.*, p.528.

¹⁶⁸ João Antunes Varela, *Ob. cit.*, p.530.

¹⁶⁹ De acordo com o preceito 823 do BGB, o sistema alemão reconhece no I. as lesões ilícitas da vida, do corpo, da saúde, da liberdade, da propriedade ou de outros direitos, e no II. menciona a violação ilícita de lei que tenha como objeto a proteção de outrem. O BGB vai ainda mais longe com o 826, ao estipular que também gera responsabilidade quando são violados os bons costumes, desde eu haja culpa. Emilio Eiranova Encinas, “Código Civil Alemán – Comentado BGB”, Marcial Pons, Madrid, 1998, pp.261 e 262 e Luís Menezes Leitão, *Ob. cit.*, p.283.

¹⁷⁰ Luís Menezes Leitão, *Ob. cit.*, p.283/284.

¹⁷¹ Cfr. arts. 24.º e seguintes da Constituição da República Portuguesa.

¹⁷² Como o direito ao nome, ao pseudónimo, direito à imagem (presentes nos artigos 70.º e seguintes do CC). Embora o caráter subjetivo de muitos destes direitos seja duvidoso, a violação destes dá também lugar à obrigação de indemnização. João Antunes Varela, *Ob. cit.*, p.534.

¹⁷³ Ou seja, dão origem a responsabilidade civil a violação dos direitos dos cônjuges à meação dos bens comuns, direitos de administração sobre os bens dos menores, entre outros. Não havendo, portanto, tutela desta responsabilidade civil delitual para os direitos familiares de natureza pessoal como o direito à fidelidade do cônjuge ou a responsabilidade parental sobre os filhos. Luís Menezes Leitão, *Ob. cit.*, p.287 e Mário Júlio de Almeida Costa, *Ob. cit.*, p.562.

utilidades proporcionadas por esse direito”, ou seja, não é uma tutela genérica ou única do património do lesado, mas sim das “utilidades” que o direito violado lhe proporcionava ¹⁷⁴. A segunda modalidade delimitada pelo artigo 483.º diz respeito à violação de disposições legais que visam tutelar interesses alheios. Estamos a falar de regras jurídicas que mesmo sem conferirem direitos subjetivos, protegem interesses particulares. Encontram-se aqui, por exemplo, a violação ao Código da Estrada, ou de leis que imponham certas providências sanitárias ou a falsificação de documentos, entre outros. Contudo, estas “normas de proteção” exigem a verificação de três pressupostos: primeiro que “à lesão dos interesses do particular corresponda a violação de uma norma legal”, de seguida têm de estar em causa “interesses alheios legítimos ou juridicamente protegidos por essa norma”, não sendo apenas interesses reflexos, mas antes os fins visados por ela e, por fim, que o dano se verifique no próprio bem jurídico que a norma em causa pretende tutelar ¹⁷⁵.

Podemos ainda apontar algumas formas especiais de ilicitude que o Código Civil estabelece, como o abuso de direito (art. 334.º), a não cedência recíproca em caso de conflito de direitos (art. 335.º), a ofensa do crédito ou do bom nome (art. 484.º) e a prestação de conselhos, recomendações e informações (art. 485.º).

Desta forma, é considerada fundamental a ilicitude do facto voluntário praticado pelo agente para que haja responsabilidade civil extracontratual por factos ilícitos, ou seja, a violação ilícita tanto de direitos subjetivos de outrem como de normas de proteção que tutelam interesses particulares. Contudo, a lei admite algumas situações específicas que podem justificar o facto do agente, afastando assim a ilicitude da sua conduta. Estas causas da exclusão da ilicitude traduzem-se no exercício de um direito, no cumprimento de um dever, na legítima defesa (art. 337.º), na ação direta (art. 336.º), no estado de necessidade (art. 339.º) e, por fim, no consentimento do lesado (art. 340.º) ¹⁷⁶.

O terceiro requisito da responsabilidade civil delitual é a culpa. Na regra geral da responsabilidade subjetiva, é necessário que o agente tenha atuado com “*dolo ou mera culpa*”, sendo apenas em casos excecionais que se admite a obrigação de indemnizar independentemente da culpa. Não basta que o comportamento do agente seja ilícito, ou seja, censurável objetivamente, é também indispensável que a sua atitude seja reprovável na medida em que, atendendo às circunstâncias concretas do caso, ele “*podia e devia de ter*

¹⁷⁴ Luís Menezes Leitão, *Ob. cit.*, p.286.

¹⁷⁵ João Antunes Varela, *Ob. cit.*, p.540.

¹⁷⁶ Luís Menezes Leitão, *Ob. cit.*, p.299.

agido de outro modo”¹⁷⁷ ¹⁷⁸. Para este efeito, o primeiro passo é saber quem é que está suscetível a este tipo de juízo de censura, ou seja, definir o conceito de imputabilidade. De acordo com a letra da lei, a imputabilidade “*equivale à capacidade natural*”, sendo, portanto, a aptidão de certa pessoa de “*entender ou querer*” os atos e as consequências ou efeitos que deles podem advir, bem como a faculdade de se determinar de acordo com os juízos por si realizados¹⁷⁹. É por isto que o n.º 1 do artigo 488.º do Código Civil refere que quem se encontrar incapacitado de entender ou querer os factos, no momento da sua ocorrência, não responde pelas consequências danosas que se verificarem, salvo se o próprio tiver contribuído culposamente para essa incapacidade e esta for transitória¹⁸⁰. O n.º 2 do mesmo preceito complementa com uma presunção de inimputabilidade para menores de sete anos.

Segundo o artigo 483.º, é possível apurar duas modalidades de culpa: o dolo e a negligência (*mera culpa*). No dolo encontramos as situações onde o agente tem a intenção de praticar o ato¹⁸¹. Por sua vez na negligência não se verifica a intenção do agente, contudo continua a ser um comportamento censurável pois aquele não demonstrou a diligência a que estava legalmente adstrito¹⁸². É uma distinção que não tem tanta relevância como no direito penal, mas que ainda assim pode influenciar a indemnização, tal como é possível observar no artigo 494.º, que permite fixar a indemnização, equitativamente, em montantes inferiores aos valores dos danos causados se a responsabilidade se fundar na mera culpa, desde que o

¹⁷⁷ João Antunes Varela, *Ob. cit.*, p.562.

¹⁷⁸ Menezes Leitão define culpa como um juízo de censura feito ao agente por ele ter adotado uma conduta contrária àquela a que estaria obrigado segundo a conduta legal. Luís Menezes Leitão, *Ob. cit.*, p.307.

¹⁷⁹ José Alberto González, “*Responsabilidade Civil*”, Quid Juris, 3.ª edição, 2013, p.164.

¹⁸⁰ Se um indivíduo, por ter ingerido substâncias psicotrópicas ou alcoólicas, ou tiver adormecido enquanto conduzia um veículo automóvel, provocar danos, responderá por eles. O seu estado de inimputabilidade foi induzido pelo próprio e, sendo transitório, responderá na medida dos seus atos. Contrariamente, caso se verifique um estado de inimputabilidade definitivo, a responsabilidade será excluída. Neste âmbito é importante mencionar o artigo 489.º/n.º 1 que, por motivos de equidade, estipula uma responsabilização parcial ou total do inimputável pelos danos, desde que não seja possível obter a devida compensação das pessoas a quem incumbe a sua vigilância (art. 491.º). Esta potencial indemnização é limitada pelo facto de não se poder privar a pessoa não imputável dos alimentos necessários, de acordo com o seu estado e condição, nem dos meios indispensáveis para cumprir os seus deveres legais de alimentos. Luís Menezes Leitão, *Ob. cit.*, p.308 e 309.

¹⁸¹ Podendo este ser um dolo direto (“*o agente quer a verificação do facto, sendo a sua conduta dirigida diretamente a produzi-lo*”), um dolo necessário (“*o agente não dirige a sua atuação diretamente a produzir a verificação do facto, mas aceita-o como consequência necessária da sua conduta*”) ou um dolo eventual (“*o agente representa a verificação como consequência possível da sua conduta e atua, conformando-se com a sua verificação*”). *Ibidem*, p.310.

¹⁸² A negligência ou *mera culpa*, pode ser consciente (“*o agente, violando o dever de diligência a que estava obrigado, representa a verificação do facto como consequência possível da sua conduta, mas atua sem se conformar com a sua verificação*”) ou inconsciente (“*o agente violando o dever de diligência a que estava obrigado, não chega a sequer representar a verificação do facto*”). *Ibidem*.

grau de culpabilidade do agente, a situação económica do lesado e lesante e as demais circunstâncias o justifiquem ¹⁸³.

Este requisito da culpa vai ser apreciado segundo o critério do homem médio, ou seja, estamos a falar de uma culpa em abstrato ¹⁸⁴. Na falta de outro critério legal, a censurabilidade da conduta do agente será avaliada pela diligência do *bonus pater familias*, um bom pai de família (“*um homem normal, medianamente sagaz, prudente, avisado e cuidadoso*” ¹⁸⁵), mas considerando sempre as circunstâncias de cada caso ¹⁸⁶. Aqui na responsabilidade delitual, a culpa carece de prova, sendo que é sobre o lesado que recai este ónus, salvo se houver uma presunção legal de culpa. Portanto, há situações específicas onde se verifica uma inversão do ónus da prova, passando esta a correr por conta do lesante. Temos assim a inversão do ónus nas situações de responsabilidade das pessoas obrigadas à vigilância de outrem (art. 491.º), nos danos causados por edifícios ou outras obras (art. 492.º), nos danos causados por coisas ou animais (art. 493.º/n.º 1), nos danos resultantes de atividades perigosas (art. 493.º/n.º 2) e no caso específico dos acidentes causados por veículos (art. 503.º/n.º 3).

No que diz respeito ao requisito da culpa, resta apenas referir a possível exclusão deste pressuposto quando o agente se encontrar em determinadas situações que afastam a possibilidade de a ordem jurídica censurar a sua conduta, como os casos de erro desculpável, medo invencível e desculpabilidade ¹⁸⁷.

O próximo requisito é o do dano. Para que o lesante esteja obrigado a indemnizar o lesado, é essencial que haja dano, ou seja, que o facto ilícito culposo tenha causado um prejuízo a alguém ¹⁸⁸. O conceito de dano, deve de ser visto, simultaneamente, num sentido fático e normativo, sendo assim a “*frustração de uma utilidade que era objeto de tutela jurídica*” ¹⁸⁹.

Várias são as classificações que se podem estabelecer relativamente ao dano, como a do dano em sentido real (correspondente à “*avaliação em abstrato das utilidades que eram*

¹⁸³ João Antunes Varela, *Ob. cit.*, p.567 e 568.

¹⁸⁴ É a teoria da culpa em abstrato, apreciada segundo o homem médio, que o código civil expressamente consagra para a responsabilidade civil delitual, de acordo com o art. 487.º n.º 2, sendo também esta mesma teoria aplicada para a responsabilidade contratual, segundo o art. 799.º/n.º 2. Contrariamente à doutrina anterior que defendia a aplicação da culpa em concreto para a responsabilidade civil contratual, ou seja, uma apreciação da culpa de acordo com a diligência que o agente habitualmente era capaz. João Antunes Varela, *Ob. cit.*, p.575; e Luís Menezes Leitão, *Ob. cit.*, p.314 e 315.

¹⁸⁵ João Antunes Varela, *Ob. cit.*, p.568.

¹⁸⁶ Cfr. art. 487.º/n.º 2.

¹⁸⁷ Luís Menezes Leitão, *Ob. cit.*, p.325 e 326.

¹⁸⁸ João Antunes Varela, *Ob. cit.*, p.597.

¹⁸⁹ Luís Menezes Leitão, *Ob. cit.*, p.327.

objeto de tutela jurídica”, implicando uma indemnização através da restituição natural do bem (reparação do objeto lesado ¹⁹⁰) ou de uma indemnização específica (entrega de um bem equivalente) ¹⁹¹) e a do dano em sentido patrimonial (correspondente a uma avaliação concreta dos efeitos da lesão no património do lesado, traduzindo-se a indemnização numa compensação pela diminuição que aquele património sofreu ¹⁹²). Dentro do dano patrimonial, encontramos ainda os danos emergentes (danos causados nos bens ou direitos já existentes na titularidade do lesado à data da lesão ¹⁹³) e os lucros cessantes (os benefícios que o lesado iria adquirir, se não fosse aquela lesão ¹⁹⁴).

Uma outra importante distinção é a dos danos patrimoniais e não patrimoniais. Os primeiros dirigem-se a uma vertente mais material, são aqueles prejuízos que correspondem “à frustração de utilidades suscetíveis de avaliação pecuniária”. Como por exemplo, a destruição de bens pertencentes ao património do lesado. Os danos não patrimoniais, ou morais, são aqueles que não são suscetíveis de avaliação pecuniária ¹⁹⁵. Estamos a falar de danos como o desgosto, o sofrimento físico, perdas de prestígio e reputação, ou seja, atingem bens que não integram o património do lesado (a saúde, o bem-estar, a liberdade, a beleza, a honra, o bom nome, entre outros) ^{196 197}. Há ainda espaço para o dano morte, de acordo com o n.º 2, 3 e 4 do artigo 496.º.

Por fim, como último pressuposto temos onexo causal entre o dano e o facto ilícito culposo. O lesante “está obrigado a indemnizar o lesado pelos danos resultantes da violação”. Várias foram as teorias que surgiram para averiguar a verificação deste pressuposto, ou seja, para confirmar que existe umnexo de causalidade entre a conduta do agente e os resultados danosos, contudo, muitas dessas teorias provaram-se insuficientes ou desadequadas, entre elas a teoria da equivalência das condições, a teoria da última condição, teoria da condição eficiente e a teoria do fim da norma violada ¹⁹⁸.

¹⁹⁰ Cfr. art. 562.º do CC.

¹⁹¹ *Ibidem*, p.328.

¹⁹² *Ibidem*.

¹⁹³ João Antunes Varela, *Ob. cit.*, p.599.

¹⁹⁴ Luís Menezes Leitão, *Ob. cit.*, p.329.

¹⁹⁵ Luís Menezes Leitão, *Ob. cit.*, p.330 e 331.

¹⁹⁶ João Antunes Varela, *Ob. cit.*, p.601.

¹⁹⁷ Cfr. art. 496.º/n.º 1, que estipula que se deve de atender aos danos não patrimoniais, que pela sua gravidade, mereçam tutela do direito.

¹⁹⁸ A teoria da equivalência das condições (ou teoria da *conditio sine qua non*) estipula que o “prejuízo deveria de ser considerado como provocado por quaisquer eventos cuja não verificação tivesse acarretado a inexistência do dano. Isto é: o nexo causal dar-se-ia a favor de qualquer evento que fosse condição necessária do dano”. (António Menezes Cordeiro, “Tratado de Direito Civil VIII”, 2.º edição revista e atualizada, Almedina, 2023, p.563). Por sua vez, a teoria da última condição considera como causa do dano o último evento verificado, ou seja, aquele que lhe antecedeu diretamente. (António Menezes Cordeiro, *ob. cit.*, p.563). Quanto à teoria da condição eficiente, o nexocausal verificar-se-á de acordo com a condição que, de entre todas as

Para este efeito a teoria adotada pela maioria da doutrina, subjacente ao artigo 563.º do Código Civil é a teoria da causalidade adequada ou teoria da adequação, de acordo com a qual, “A obrigação de indemnizar só existe em relação aos danos que o lesado provavelmente não teria sofrido se não fosse a lesão”. De acordo com a jurisprudência, a ideia fulcral “*é a de que se considera causa de um dano o facto que se revele, em concreto, condição necessária desse dano, mas também que constitua, em abstrato, segundo o curso normal das condições, causa adequada da sua produção*”¹⁹⁹. Segundo Almeida Costa, “*considera-se causa de um prejuízo a condição que, em abstrato, se mostra adequada a produzi-lo*”²⁰⁰. A averiguação desta adequação, será realizada *a posteriori*, através de um juízo de prognose póstuma, tomando-se por base “*não apenas as circunstâncias normais que levariam um observador externo a efetuar um juízo de previsibilidade, mas também as circunstâncias anormais, desde que reconhecíveis ou conhecidas pelo agente*”²⁰¹.

3.2.2. Responsabilidade Civil pelo Risco

A responsabilidade civil pelo risco ou responsabilidade objetiva, tal como foi dito anteriormente e em concordância com o n.º 2 do artigo 483.º, é uma possibilidade excepcional, permitida apenas nos casos especificados na lei. Portanto, pode haver obrigação de indemnizar independentemente de culpa, desde que a lei assim o preveja.

Estamos a falar de um instituto que foi particularmente inspirado pela teoria do risco, que deste modo estipula que quem exerce uma atividade considerada perigosa, responde pelos danos que ocasionalmente ocorram, mesmo que sem culpa²⁰².

Deste modo os casos onde é admissível esta responsabilidade objetiva serão situações como a atuação das pessoas em proveito alheio (arts. 500.º e 501.º), a utilização de coisas perigosas, como animais, veículos e instalações de energia elétrica e gás (arts. 502.º, 503.º e 509.º), podendo ainda se encontrar responsabilidade pelo risco em situações constantes em

condições do processo causal, apresente uma maior eficácia de produzir o dano (Luís Menezes Leitão, *Ob. cit.*, p.341). Por fim, a teoria do fim na norma violada considera que para se averiguar o nexo de causalidade tem de se apurar se o dano resultante desse facto corresponde ao fracasso das utilidades que a norma violada visa tutelar (Luís Menezes Leitão, *Ob. cit.*, p.341)

¹⁹⁹ Ac. Tribunal da Relação do Porto de 26/06/2023, Processo n.º 1483/21.0T8VNG.P1, relator Miguel Baldaia de Morais.

²⁰⁰ Mário Júlio de Almeida Costa, *Ob. cit.*, p.763.

²⁰¹ Ac. Tribunal da Relação do Porto de 26/06/2023, Processo n.º 1483/21.0T8VNG.P1, relator Miguel Baldaia de Morais.

²⁰² Mário Júlio de Almeida Costa, *Ob. cit.*, p.613.

diplomas especiais, como é o caso da responsabilidade do produtor por produtos defeituosos
203 204 .

3.2.2.1. A responsabilidade do Comitente

O artigo 500.º/n.º 1 estipula o que “*Aquele que encarrega outrem de qualquer comissão responde, independentemente de culpa, pelos danos que o comissário causar, desde que sobre este recaia também a obrigação de indemnizar.*” O n.º 2 complementa, referindo que “*A responsabilidade do comitente só existe se o facto danoso for praticado pelo comissário, ainda que intencionalmente ou contra as instruções daquele, no exercício da função que lhe foi confiada.*”.

Esta hipótese implica a responsabilização de alguém que encarregou outro sujeito de determinada função, quando este tenha cometido um facto danoso no exercício das funções que lhe foram atribuídas, mesmo que o primeiro, comitente, não tenha qualquer culpa. Ainda assim, o comitente vai gozar de um direito de regresso face ao comissário, tal como dita o n.º 3 do artigo 500.º, “*O comitente que satisfizer a indemnização tem o direito de exigir do comissário o reembolso de tudo quanto haja pago*”, salvo se aquele também tiver procedido com culpa, caso em que o direito se irá moldar à medida da culpa de cada um ²⁰⁵. Esta relação externa (comitente-lesado) e interna (comitente-comissário) tem a função de garantir que o lesado é indemnizado pelos danos sofridos já que, normalmente, o comissário apresenta um património pouco suscetível de suportar o pagamento de elevadas indemnizações, ao contrário do comitente, sendo assim uma hipótese mais justa ²⁰⁶.

Portanto, para que este regime de responsabilidade objetiva do comitente pelos atos danosos do comissário seja aplicado, é necessário que se verifiquem três pressupostos: primeiro, a existência de uma relação de comissão, depois que os factos danosos praticados pelo comissário tenham ocorrido no exercício da função que lhe foi confiada, e por último, que também sobre o comissário recaia a obrigação de indemnizar.

No que diz respeito à relação de comissão, aqui está em causa o seu sentido amplo, sendo por isso vista como um serviço ou função realizada no interesse e por conta de outrem (comitente), podendo ser uma atividade duradoura ou um ato isolado, atos materiais ou

²⁰³ Regulamentado ao abrigo do Decreto-Lei n.º 383/89, de 6 de novembro.

²⁰⁴ Luís Menezes Leitão, *Ob. cit.*, p.363.

²⁰⁵ Cfr. art.497.º/n.º 2, *ex vi* art.500.º/n.º 3 *in fine*.

²⁰⁶ João Antunes Varela, *Ob. cit.*, p.646.

jurídicos, de carácter gratuito ou oneroso, entre outros ²⁰⁷. Uma boa parte da doutrina defende ainda a exigência de características como a existência de um nexo de subordinação do comissário ao comitente, onde este dá ordens e instruções àquele para que assim esteja justificada a possibilidade de um responder pelos atos do outro ²⁰⁸, contudo, parece fazer mais sentido a visão contrária de Menezes Leitão que, recorrendo ao n.º 2 do artigo 500.º, expressamente refere que o comitente vai responder mesmo que o comissário atue “*contra as instruções daquele*” ²⁰⁹. O mesmo acontece para a exigência da liberdade de escolha do comitente sobre o comissário, que também pode ser afastado porque a lei estabelece um regime de “*responsabilidade objetiva de garantia da indemnização*” e não uma responsabilidade do comitente por *culpa in eligendo* ²¹⁰. Por fim, é relevante referir que esta relação de comissão não engloba todo e qualquer serviço, mas apenas aqueles que possam ser imputados ao comitente, ou seja, “*são praticados exclusivamente no seu interesse*” sendo ele que suporta as “*despesas e ganhos dessa atividade*” ²¹¹.

Como segundo pressuposto temos a necessidade de o comissário praticar os factos danosos ao abrigo das funções que lhe foram confiadas pelo comitente, portanto a responsabilidade deste não deve “*extravasar da função que lhe foi efetivamente confiada, funcionando esta assim como delimitação da zona de riscos a cargo do comitente*” ²¹². Também aqui alguma doutrina adotou uma visão que levanta alguma querela. Está aqui em causa uma interpretação restritiva que defende que os atos que não se inseriam no esquema do exercício da função ou que teriam um fim ou interesse que lhe fosse estranho, não seriam abrangidos por esta responsabilidade a não ser que houvesse um nexo instrumental entre a função e os danos ²¹³. Voltamos a concordar com Menezes Leitão que, discordando desta visão, afirma que esta interpretação “*retiraria grande parte do alcance à responsabilidade do comitente, e não tem suporte legal.*” Assim basta apenas um “*nexo etiológico entre a função e os danos, no sentido de que seja no seu exercício que os danos sejam originados*” ²¹⁴.

²⁰⁷ Luís Menezes Leitão, *Ob. cit.*, p.364 e João Antunes Varela, *Ob. cit.*, p.640.

²⁰⁸ João Antunes Varela, *Ob. cit.*, p.640 e 641.

²⁰⁹ Luís Menezes Leitão, *Ob. cit.*, p.364/365.

²¹⁰ *Ibidem*, p.365.

²¹¹ Enquadram-se nestas relações de comissão por exemplo, os contratos de trabalho e contratos de mandato. Não englobando as situações em que quem realiza o serviço a pedido exerce uma função autónoma que não lhe foi delegada pelo comitente, como nos contratos de transporte, de empreitada, ou no depósito. *Ibidem*;

²¹² *Ibidem*, p.366.

²¹³ Mário Júlio de Almeida Costa, *Ob. cit.*, p.619 e João Antunes Varela, *Ob. cit.*, p.646.

²¹⁴ Luís Menezes Leitão, *Ob. cit.*, p.366.

Por fim, exige-se também que sobre o comissário recaia a obrigação de indemnizar. A divergência doutrinária aqui decorre sobre a exigência ou não da culpa do comissário. Antunes Varela diz-nos que “*o comitente só responde (objetivamente) quando haja culpa do comissário*”²¹⁵. Menezes Leitão admite as dúvidas levantadas pela interpretação da lei referindo que o n.º 1 do artigo 500.º em nenhum momento aparenta referir a culpa, por outro lado, aquando do n.º 3, que estabelece limitações ao direito de regresso, há uma menção expressa a “*também tiver culpa*”, na eventualidade de o comitente agir culposamente. Acaba por se decidir pela mesma tese, “*parece que a lei não exigirá uma demonstração efetiva da culpa do comissário, bastando-se o art.500.º, n.º 1, com uma culpa presumida*”²¹⁶.

3.2.2.2. A responsabilidade do produtor por produtos defeituosos

Um outro regime igualmente importante é o da responsabilidade do produtor por produtos defeituosos presente no Decreto-Lei n.º 383/89, de 6 de novembro, que transpõe para a ordem jurídica interna a Diretiva n.º 85/374/CEE do Conselho de 25 de julho de 1985²¹⁷.

O princípio fundamental que está aqui em causa é o da responsabilidade objetiva do produtor, ou seja, uma responsabilidade independentemente de culpa pelos danos causados por defeitos dos produtos que aquele põe em circulação²¹⁸. Estamos assim inseridos no âmbito da responsabilidade extracontratual, o que significa que este regime depende da verificação dos pressupostos anteriormente mencionados, com a exceção do requisito da culpa. O lesado terá então de provar o “*dano, o defeito e o nexo de causalidade entre aquele e este para ter direito a ser indemnizado*”²¹⁹.

De acordo com esta legislação, é o produtor que vai responder objetivamente pelos produtos que colocar em circulação, como tal temos de identificar o que estará abrangido pela categoria de “produtor”. Para este efeito, no n.º 1 e n.º 2 do artigo 2.º do Decreto-Lei n.º 383/89, o legislador apresenta-nos uma definição um tanto ampla, considerando como produtor o “*fabricante do produto acabado, de uma parte componente ou de matéria-prima, e ainda quem se apresente como tal pela aposição no produto do seu nome, marca ou outro*

²¹⁵ João Antunes Varela, *Ob. cit.*, p.644.

²¹⁶ Luís Menezes Leitão, *Ob. cit.*, p.368 e 369.

²¹⁷ Quanto às alterações sofridas e propostas, será, posteriormente, dedicada a respetiva referência e análise.

²¹⁸ Cf. art. 1.º do Decreto-Lei n.º 383/89, de 06 de novembro; João Calvão da Silva, “*Responsabilidade Civil do Produtor*”, Almedina, Coimbra, 1990, p.487.

²¹⁹ João Calvão da Silva, *Ob. cit.*, p.490.

sinal distintivo.” Mas o legislador não se ficou por aqui, também considerou como produtor “*aquele que na Comunidade Económica Europeia e no exercício da sua atividade comercial, importe do exterior da mesma produtos para venda, aluguer, locação financeira ou outra qualquer forma de distribuição*”, e ainda “*qualquer fornecedor de produto cujo produtor comunitário ou importador não esteja identificado, salvo se notificado por escrito, comunicar ao lesado no prazo de três meses, igualmente por escrito, a identidade de um ou outro, ou a de algum fornecedor precedente*”²²⁰.

A aplicação desta responsabilidade pelo risco exige ainda um defeito no produto. Em primeiro lugar, no que ao produto diz respeito, o artigo 3.º entende-o como qualquer coisa móvel, ainda que incorporada noutra coisa móvel ou imóvel. Quanto ao defeito, considera-se defeituoso o produto que “*não oferece a segurança com que legitimamente se pode contar, tendo em conta as circunstâncias, designadamente a sua apresentação, a utilização que dele razoavelmente possa ser feita e o momento da sua entrada em circulação*”, não se considerando defeituoso um produto apenas pelo facto de posteriormente entrar em circulação um outro mais aperfeiçoado^{221 222}. Para averiguar a segurança com que legitimamente se pode contar, o critério a que se deve de atender não são as expectativas subjetivas do consumidor lesado, mas sim as expectativas objetivas do público em geral, ou

²²⁰ É uma definição com bastante amplitude considerando uma divisão tripartida da definição de produtor. Temos o produtor real ou em sentido estrito, como o fabricante do produto acabado, de uma parte componente ou de matéria-prima, ou seja, qualquer pessoa humana ou pessoa jurídica que participe na criação de um destes elementos responderá pelos defeitos que neles se apresentem. João Calvão da Silva, *Ob. cit.*, p.546. Também incorpora o produtor aparente, que será aquele que se apresente como tal pela aposição no produto do seu nome, marca ou outro sinal distintivo. Encontram-se aqui os grandes distribuidores, as cadeias comerciais e as empresas de vendas por correspondência. João Calvão da Silva, *Ob. cit.*, p.550; Por fim temos também o produtor presumido, que engloba as situações previstas no n.º 2 do artigo 2.º, ou seja, aquele que importa do exterior da Comunidade Económica Europeia, produtos para venda, aluguer, locação financeira ou qualquer outra forma de distribuição, no exercício da sua atividade comercial, bem como o fornecedor de produtos onde o produtor comunitário ou importador não estão identificados, quando aquele não tenha comunicado as informações necessárias à identificação destes últimos, ao lesado. João Calvão da Silva, *Ob. cit.*, p.554.

²²¹ Cf. art. 4.º/ n.º 1 e n.º 2.

²²² Quanto à “segurança com que legitimamente se pode contar”, o núcleo da questão é a segurança do produto e não a idoneidade deste para a realização do fim a que se destina. Contrariamente ao que acontece na responsabilidade contratual, onde o vício traduz-se na falta de conformidade ou qualidade das coisas (visa proteger “*o interesse subjacente ao cumprimento perfeito*”), aqui, esta “moderna responsabilidade do produtor”, caracteriza-se “*justamente por ser uma responsabilidade por falta de segurança dos produtos*” (protegendo “*a integridade pessoal do consumidor e dos seus bens*”) João Calvão da Silva, *Ob. cit.*, p.634 e 635; Vera Lúcia Coelho, “Responsabilidade do produtor por produtos defeituosos. Tese de resistência ao DL n.º 383/89, de 6 de novembro, à luz da jurisprudência recente, 25 anos volvidos sobre a sua entrada em vigor”, em *RED - Revista Eletrónica de Direito*, n.º 2, junho, 2017.

seja, a “segurança esperada e tida por normal nas conceções do tráfico do respetivo setor de consumo”²²³.

É pertinente referir que o legislador optou por não distinguir expressamente a tipologia dos defeitos, ficando-se por um regime mais uniforme de responsabilidade objetiva que engloba os vários tipos. Ainda assim, podemos elencar os seguintes tipos de defeitos: defeitos de conceção, defeitos de fabrico e defeitos de informação. Primeiramente, os defeitos de conceção, ocorrem no plano da idealização do produto, ou seja, surgem deficiências ou erros logo na fase inicial de planeamento e preparação, traduzindo-se assim os vícios em todos os produtos da série ou séries fabricadas. Por sua vez, os defeitos de fabrico, são os que surgem na “fase de laboração, produção ou fabrico”, sendo resultado de um *design* ou projeto perfeito, sem quaisquer vícios. Ocorrem, usualmente, devido a falhas humanas ou mecânicas em indústrias de produção em massa, automatizada e estandardizada²²⁴. Os defeitos de informação são os devidos por “falta, insuficiência ou inadequação das informações, advertências ou instruções sobre o seu uso e perigos conexos”²²⁵. A falta deste tipo de informações pode levar ao mau uso ou emprego de produtos em causa, culminando em ainda mais danos²²⁶.

É ainda considerada uma quarta categoria de defeitos, os defeitos de desenvolvimento onde “um produto pode ser ilegitimamente inseguro por riscos ou defeitos incognoscíveis perante o estado da ciência e da técnica existente ao tempo da sua emissão no comércio”²²⁷. São riscos que se podem encontrar no artigo 5.º do Decreto-Lei, como uma das causas que leva à exclusão da responsabilidade, precisamente por serem defeitos indetetáveis pelo produtor quando coloca os produtos em circulação, permanecendo escondidos até que os progressos científicos e técnicos permitam a sua deteção²²⁸.

De acordo com o artigo 5.º, o produtor verá a sua responsabilidade excluída quando se verificarem algumas das circunstâncias previstas nas alíneas deste preceito²²⁹, como por exemplo se ele provar que não colocou, “voluntariamente”, o produto em circulação (alínea a)), ou provar que está em causa um risco de desenvolvimento, onde o estado do

²²³ É o juiz que, atendendo à peculiaridade do produto em causa e todas as circunstâncias do caso concreto, determinará se um produto é defeituoso dado o grau de insegurança conforme as legítimas expectativas do público em geral ou não. João Calvão da Silva, *Ob. cit.*, p.636 e 637.

²²⁴ João Calvão da Silva, *Ob. cit.*, p.655.

²²⁵ *Ibidem*, p.658.

²²⁶ Vera Lúcia Coelho, *Ob. cit.*, p.25.

²²⁷ João Calvão da Silva, *Ob. cit.*, p.662.

²²⁸ Vera Lúcia Coelho, *Ob. cit.*, p.26.

²²⁹ João Calvão da Silva, *Ob. cit.*, pp.717 e seguintes.

conhecimento científico e técnico não permitiam detetar a existência do defeito, no momento em que o produto foi colocado em circulação (alínea e) ²³⁰, entre outros.

Quanto aos danos ressarcíveis, segundo o artigo 8.º, nos casos de morte ou lesão pessoal, todos os danos patrimoniais e não patrimoniais são indemnizáveis ²³¹. Também são ressarcíveis os danos verificados em coisa diversa do produto defeituoso, desde que seja normalmente destinado ao uso ou consumo privado e o lesado lhe tenha dado principalmente este destino. Contudo, o artigo 9.º impõe uma limitação para os danos causados em coisas, pois só são indemnizáveis a partir do montante de 500 (euros) ²³².

Na eventualidade de existirem várias pessoas responsáveis pelos danos, a responsabilidade será solidária, nos termos do artigo 6.º/ n.º 1. Havendo um facto culposo do lesado que tenha concorrido para o dano, o tribunal poderá reduzir ou excluir a indemnização, depois de consideradas todas as circunstâncias ²³³.

Para terminar, é importante referir o nexo de causalidade, que é um dos elementos constitutivos fundamentais da responsabilidade civil. O produtor responde, independentemente de culpa, pelos danos causados pelos defeitos do produto que colocou em circulação. “*Nem todos e quaisquer danos sobrevindos ao defeito do produto são incluídos na responsabilidade do produtor*” ²³⁴. Estamos a falar de umas das três provas que incumbem ao lesado (o defeito, o dano e o nexo causal entre defeito e dano), algo que se afigura de enorme dificuldade ²³⁵. Na apreciação da prova “*valem as regras do direito comum*”, referindo Calvão Silva que o lesado deve de ser auxiliado na difícil tarefa de demonstrar o nexo de causalidade “*através de uma prova de primeira aparência*”. Com o defeito e o dano provados, pode haver uma espécie de presunção da causalidade através do uso das regras da “*experiência da vida e teoria da causalidade (que reconduz o nexo causal a um juízo de probabilidade)*” ²³⁶.

²³⁰ A presença dos riscos de desenvolvimento como uma das causas de isenção da responsabilidade do produtor, suscita algumas críticas, principalmente no que diz respeito à natureza da responsabilidade presente neste regime. Ao “*excluir os danos resultantes dos riscos de desenvolvimento, aproxima bastante esta responsabilidade de uma responsabilidade por culpa*.” (Luís Menezes Leitão, *Ob. cit.*, p.396). “*Não equivalerá a aceitação de tal causa de exclusão da responsabilidade à admissão da prova da inexistência de culpa por parte do produtor?*” (Sobre esta questão leia-se João Calvão da Silva, *Ob. cit.*, p.503).

²³¹ João Calvão da Silva, *Ob. cit.*, p.678.

²³² Contrariamente ao regime anterior que estipulava uma limitação máxima de indemnização para os casos de morte ou lesão pessoal (em cerca de 10 milhões de escudos). Mudança esta criticada por Menezes Leitão pois ao “*estabelecer uma responsabilidade objetiva ilimitada nestes momentos só poderá servir para agravar a insegurança jurídica dos produtores, retraindo a atividade económica*.” Luís Menezes Leitão, *Ob. cit.*, p.397, nota 910.

²³³ Cf. artigos 6.º e 7.º.

²³⁴ João Calvão da Silva, *Ob. cit.*, p.711.

²³⁵ Vera Lúcia Coelho, *Ob. cit.*, p.35.

²³⁶ João Calvão da Silva, *Ob. cit.*, p.712 e 713.

3.3. Responsabilidade Civil pelos danos causados por Sistemas de Inteligência Artificial

Na construção atual da sociedade já não é de estranhar a presença assídua de todo um conjunto de dispositivos tecnológicos que automatizam não só as tarefas relacionadas com produção industrial em massa, associada a trabalhos mais complexos e pesados, mas também inúmeras atividades de menor calibre consideradas redundantes, isto sem nunca esquecer a sociedade digital na qual hoje nos inserimos. Foi este célere crescimento da sociedade automatizada, quase que de forma inconspícua, que nos trouxe ao debate em questão.

Sistemas de inteligência artificial presentes nos dispositivos móveis e computadores, presentes nos automóveis (veículos autónomos), nas nossas casas (*smart homes* ou casas inteligentes) e em seus sistemas de segurança, presentes nos sistemas de funcionamento geral de indústrias, institutos e organizações, seja através de *hardware* humanoide ou não (*robots* ou só *software*), implicam uma rigorosa ponderação na apresentação de soluções associadas às questões de responsabilidade civil por eventuais danos que venham a ser observados.

Acontece que, os regimes tradicionais de responsabilidade civil não são suficientes nem os mais eficazes na proteção do lesado por eventuais danos provocados por entes dotados de inteligência artificial. Não se trata de uma inaplicabilidade dos regimes existentes, mas sim da possibilidade de deixar a vítima pouco ou nada compensada, dados os desafios que se colocam e as limitações existentes dos sistemas de responsabilidade clássicos, dificuldade esta que é sentida pelas características particulares que os sistemas de inteligência artificial apresentam como a autonomia e a imprevisibilidade (provocada pelas técnicas de *machine learning* e *deep learning*), a conectividade, a complexidade, a opacidade dos processos de decisão e a dependência de dados²³⁷.

Várias são as questões que se colocam em torno da IA, mas os problemas centrais que mais tem importunado os juristas estão diretamente relacionados com a responsabilidade subjetiva e os seus requisitos cumulativos, em especial a culpa e o nexo de causalidade.

²³⁷ Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.19; e Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, p.243.

Comecemos com o problema da culpa que é uma das maiores incógnitas que aqui se coloca. Tal como referido anteriormente o pressuposto da culpa é fundamental para que nos termos do artigo 483.º do Código Civil, recaia sobre o lesante a obrigação de indemnizar. Esta ideia é reforçada pela atribuição do carater excecional à responsabilidade objetiva, ou seja, só nos casos especificados na lei é que teremos responsabilidade do lesante independentemente da culpa. Esta necessidade de um juízo de censura que recaí sobre a conduta do agente vai suscitar algumas dúvidas.

Num primeiro momento, temos a questão da identificação do lesante a quem imputar esse juízo de reprovação da conduta, algo que não se afigura tarefa fácil pois as “*novas tecnologias aproximam os utilizadores, mas diluem as identidades dos agentes responsáveis*”²³⁸. Quem vai responder pelos danos causados por estes sistemas de inteligência artificial? A própria máquina? A pessoa singular ou coletiva que dela tira benefício? O produtor ou fabricante, quer do *software* quer do *hardware*? A resposta mais direta e arrojada seria a responsabilização do próprio sistema tecnológico, algo que não se demonstra, por enquanto, viável. A este propósito entramos em contacto com outra grande temática que envolve o direito e a inteligência artificial que consiste na atribuição de personalidade jurídica a estes entes tecnológicos. Terá uma máquina, um robô autónomo, um algoritmo inteligente, personalidade jurídica? Poderá uma IA ser titular de direitos e obrigações? O próprio Parlamento Europeu deu início a este debate, solicitando à Comissão uma ponderação no sentido da eventual criação de um “*estatuto jurídico específico para os robôs a longo prazo, de modo que, pelo menos, os robôs autónomos mais sofisticados possam ser determinados como detentores do estatuto de pessoas eletrónicas responsáveis por sanar quaisquer danos que possam causar e, eventualmente, aplicar a personalidade eletrónica a casos em que os robôs tomem decisões autónomas ou em que interagem por qualquer outro modo com terceiros de forma independente*”²³⁹. Esta solução parece, no momento, inconcebível. De acordo com Mafalda Miranda Barbosa, um algoritmo inteligente ser titular de direitos e obrigações jurídicas pressupõe uma capacidade de entendimento e consciência que, no momento, parece difícil de se atingir artificialmente, “*falta-lhes, em cada tomada de decisão, a pressuposição ética, falha a relação de cuidado com o outro (...)*”, chegando mesmo a autora a considerar “*desdignificante para o ser humano*” a

²³⁸ Henrique Sousa Antunes, “Inteligência Artificial e responsabilidade civil: Enquadramento,” em *Revista de Direito da Responsabilidade*, ano 1, 2019, p.1/2.

²³⁹ Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de direito civil sobre robótica (2015/2103(INL)), ponto 59, alínea f).

comparação deste com um robô ou outro mecanismo de inteligência artificial^{240 241}. Neste sentido, e tendo em consideração o nível atual da tecnologia, partilhamos esta posição que recusa a extensão da personalidade jurídica aos sistemas de inteligência artificial reconhecendo, contudo, que é importante estarmos cientes de que este debate está longe do seu fim e pode sofrer alterações significativas à luz do progresso científico e tecnológico²⁴²
243.

Desta forma, não encontramos argumentos fortes o suficiente que sustentem a atribuição de personalidade jurídica à máquina, maioritariamente pela falta de consciência no momento da decisão, falta-lhe “*uma consciência de si próprio e da relação com os outros*”, impedindo assim a eventual imputação de um juízo de censura para que ela seja responsável pelos danos causados²⁴⁴. “*Perante esta realidade, é incoerente e diríamos impossível, imputar ao robot um juízo de culpa e por conseguinte, responsabilizá-lo, quando não possui qualquer raciocínio ético, entendimento suficiente para nortear as suas condutas, conceitos de bem e de mal, de ordem jurídica, etc*”²⁴⁵. Tal como refere José González, “*Uma máquina não é, para já capaz de tanto*”, contudo, prevê-se que “*com*

²⁴⁰ Mafalda Miranda Barbosa, “Inteligência Artificial, E-Persons e Direito: Desafios e Perspetivas”, em *Revista Jurídica Luso-Brasileira*, ano 3, n.º 6, 2017, p.1482; também em Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, em *Revista de Direito da Responsabilidade*, ano 2, 2020, pp. 309 e seguintes.

²⁴¹ Também Ana Rita Maia partilha da mesma posição quanto ao não reconhecimento da personalidade jurídica à máquina, “*ainda que inicialmente se cogitasse a possibilidade de lhes estabelecer um regime próximo da personalidade, não são considerados (para já) nem pessoas jurídicas, nem (jamais) pessoas biológicas.*” Ana Rita Maia, “A Responsabilidade Civil na Era da Inteligência Artificial – Qual o caminho?”, em *Julgar Online*, maio, 2021, p.31.

²⁴² Posição que também foi adotada pelo Parlamento Europeu em 2020, contrariando assim a sua visão original da criação de um estatuto jurídico específico para robots. “*Observa, a esse respeito, que não é necessário conferir personalidade jurídica aos sistemas de IA*”. Resolução do Parlamento Europeu, de 20 de outubro de 2020, que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial (2020/2014(INL)), ponto 7; também o Expert Group on Liability and New Technologies recomendou o mesmo quando disse em 2019 “*the experts believe there is currently no need to give a legal personality to emerging digital technologies*” Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.37 e 38.

²⁴³ Uma forma mais interessante de abordar a questão da atribuição de personalidade jurídica pode encontrar-se por analogia à personalidade jurídica das pessoas coletivas. Ainda assim, alguns autores recusam também a atribuição de personalidade jurídica por alusão a esta via. De acordo com Mafalda Miranda Barbosa a atribuição de personalidade jurídica a entes coletivos “*trata-se de uma personalidade jurídica funcionalizada à prossecução de determinados interesses humanos coletivos ou comuns (...). É exatamente esta justificação que falha por respeito aos entes dotados de inteligência artificial. Inexiste um interesse humano que possa ser mais bem prosseguido com a atribuição do estatuto aos robots, exceto se virmos na não responsabilidade do proprietário ou utilizador a principal razão para a personificação.*” Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, p. 314; também Mafalda Miranda Barbosa, “Inteligência Artificial, E-Persons e Direito: Desafios e Perspetivas”, pp.1485 e 1486.

²⁴⁴ Henrique Sousa Antunes, *Ob. cit.*, pp.144 e 145.

²⁴⁵ Ana Rita Maia, *Ob. cit.*, pp.31 e 32.

facilidade (...) este argumento não valha por muito tempo” ²⁴⁶. Importante mencionar também a questão do património, intrinsecamente ligada ao reconhecimento da personalidade jurídica. Tal como refere o artigo 601.º a garantia geral das obrigações é o património do devedor, o que nos leva a questionar como poderia uma máquina estar “*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*” ^{247 248}.

A alternativa à responsabilização do próprio sistema, de acordo com uma resolução do Parlamento Europeu de 2020, reside na atribuição da “*responsabilidade às diferentes pessoas da cadeia de valor que criam, fazem a manutenção ou controlam os riscos associados ao sistema de IA*” ²⁴⁹. No momento, os sistemas de inteligência artificial, apesar das suas “supercapacidades”, têm sido “*simplesmente considerados como uma extensão da pessoa que se aproveita das utilidades por ela proporcionadas. O que significa, para todos os efeitos, que à máquina não se podem assacar ações ou inações (...) a máquina não passa, apenas, de uma maneira de a pessoa (que lhe dá uso) se apresentar ante outra*” ²⁵⁰.

Ainda no âmbito da culpa, é importante lembrar que de acordo com o n.º 1 do artigo 487.º, é ao lesado que incumbe a prova da culpa do autor da lesão, salvo havendo presunção legal de culpa. Ou seja, a regra geral é que é por conta do lesado que corre o ónus da prova da culpa do agente, prova esta que é de difícil realização (a chamada “*probatio diabolica*”). Este ónus, a cargo do lesado, “*reduz em grande medida as suas possibilidades efetivas de obter indemnização*” ²⁵¹. Uma dificuldade que se faz sentir ainda mais, numa área tão técnica

²⁴⁶ Querendo responsabilizar o sistema de um veículo autónomo por danos que este tenha causado, implicaria que ao próprio automóvel se imputasse a violação de um direito alheio, o que sugere a existência de deveres aos quais o veículo estaria adstrito a respeitar. “*O referido automóvel, em situações de responsabilidade aquiliana, se encontraria subordinado ao dever geral de respeito por direitos (absolutos) alheios. E que, em hipóteses de responsabilidade contratual, ele estaria vinculado pela obrigação negocial cujo não cumprimento lesasse o credor. (...) A imposição de deveres pressupõe que o sujeito por eles atingido tenha entendimento suficiente para nortear a sua conduta de acordo com as respetivas prescrições. Assenta, por outras palavras, na presunção de que ele é capaz de realizar a conduta ordenada ou proibida, que atua conscientemente quando entende não lhe dar cumprimento e que, por isso merece censura se, porventura, não executar o comportamento comandado.*” José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, p.93 e 94.

²⁴⁷ José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, p.95 e 96.

²⁴⁸ Para mais sobre o tema da criação de um estatuto de Pessoas Eletrónicas (*E-Persons*), atribuição de direitos a entes tecnológicos, que por si só é complexo e implicaria uma extensa exposição, consultar Mafalda Miranda Barbosa, “Inteligência Artificial, E-Persons e Direito: Desafios e Perspetivas”, em *Revista Jurídica Luso-Brasileira*, ano 3, n.º 6, 2017; também Nuno Sousa e Silva, “Direito e Robótica: uma primeira aproximação”, em *Revista da Ordem dos Advogados*, ano 77, vol. I/II, janeiro/junho, 2017, pp.505 e seguintes.

²⁴⁹ Resolução do Parlamento Europeu, de 20 de outubro de 2020, que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial (2020/2014(INL)), ponto 7.

²⁵⁰ José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, p.86.

²⁵¹ Luís Menezes Leitão, *Ob. cit.*, p.317/318.

como esta, que envolve um conhecimento complexo e rigoroso. Este foi um obstáculo referido no relatório do Expert Group on Liability and New Technologies, “*The more complex the circumstances leading to the victim’s harm are, the harder it is to identify relevant evidence*”. Este grupo reforçou ainda que pode ser muito difícil e dispendioso o processo de identificação de um erro (*bug*) no longo e complicado código de *software*, bem como todo o processo de investigação para apurar a obtenção de determinados resultados (descobrir o processo de “pensamento” do algoritmo, como é que a *input data* levou a determinada *output data* (resultado). Também pode ser “*difícil, muito longo e bem dispendioso*”, não estando de todo ao alcance da vítima comum ²⁵². A solução talvez passe pelo estabelecimento de uma presunção legal de culpa ou pelo afastamento deste requisito, possivelmente com recursos a regimes de responsabilidade objetiva.

A outra grande incógnita que aqui se coloca está relacionada com onexo causal que tem de se verificar entre o dano e a conduta lesiva do agente. A responsabilidade “*deverá estender-se (...) aos agentes que tão só colaborem na prática do dano ou a facilitem*” ²⁵³. Acontece que, esta área da inteligência artificial está a ficar cada vez mais imprevisível, imprevisibilidade esta que é o resultado do seu próprio desenvolvimento.

Ora, isto é um problema pois tem-se verificado um aumentar da complexidade destes sistemas de IA, fazendo com que a origem do dano, a causa que o desencadeou, se torne dispersa e mais difícil de identificar (“*funcionando o sistema de inteligência artificial como um verdadeiro ecossistema, pode tornar-se dilemática a descoberta do real causador do dano*” ²⁵⁴). A lesão poderá resultar da maneira como o sistema foi programado ou resultar do uso que lhe foi dado pelo seu operador ou utilizador. O nível de dificuldade desta questão cresce se estiver em causa um sistema artificial munido de um algoritmo de *machine learning* ou *deep learning* que lhe permite uma capacidade de aprendizagem e adaptação elevada, tornando as suas ações mais imprevisíveis de acordo com a informação que vai recolhendo e analisando ²⁵⁵. Podemos ainda mencionar eventuais questões de segurança seja nos casos em que as atualizações de segurança foram ou não realizadas, seja na possibilidade de ciberataques e invasões ilegítimas dos sistemas (*hackers* que interfiram no funcionamento

²⁵² Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.24.

²⁵³ Henrique Sousa Antunes, *Ob. cit.*, p.140.

²⁵⁴ Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, em *Revista de Direito da Responsabilidade*, ano 5, 2023, p. 338/339.

²⁵⁵ Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.20.

do software ²⁵⁶). Ou seja, “a interconectividade e a dependência de dados, associados à implicação de diversos sujeitos para que o sistema se torne operante, conduzem as mais das vezes a problemas atinentes à multiplicidade causal” ²⁵⁷.

Como Mafalda Miranda Barbosa também refere, são essas características (supramencionadas) da autonomia e autoaprendizagem dos sistemas artificiais que aumentam a opacidade da fronteira entre os danos resultantes de erro humano e aqueles devidos ao próprio algoritmo ²⁵⁸. Estes comportamentos imprevisíveis, diminuem o poder de controlo humano sobre estes entes tecnológicos (algoritmos que decidem “*por si como agir, que se desenvolvem, como resultado de um deep learning, sem controlo humano, torna impossível conezionar um eventual dano que possa eclodir com uma conduta negligente do ser humano*” ²⁵⁹) tornando imperativo que se encontre uma solução apropriada.

Assim, conseguimos concluir que os regimes de responsabilidade assentes na culpa, ou seja, os regimes de responsabilidade subjetiva não são os mais eficazes na tutela dos lesados por sistemas de inteligência artificial, principalmente devido à dificuldade de verificação dos requisitos legais essenciais para a sua aplicação, sobretudo o requisito da culpa e do nexo causal ²⁶⁰. Este é um desafio que tem origem nas características únicas e específicas destas novas tecnologias, tal como já tivemos oportunidade de mencionar atrás ²⁶¹.

Por estas razões, parece que a alternativa passa pela aplicação de um regime de responsabilidade objetiva (“*Strick Liability*”), que sendo um modelo que prescinde do elemento subjetivo da culpa, parece ser “a forma mais eficiente de lidar com os danos causados por agentes com IA” ²⁶². Estamos a falar de um modelo que, assente na teoria do risco, faz recair sobre o lesante a obrigação de indemnizar independentemente da culpa, ou

²⁵⁶ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, p. 284.

²⁵⁷ Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, pp. 338 e 339.

²⁵⁸ Mafalda Miranda Barbosa, “Inteligência Artificial, responsabilidade civil e causalidade: breves notas”, em *Revista de Direito da Responsabilidade*, ano 3, 2021, p. 606.

²⁵⁹ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 284.

²⁶⁰ Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, pp.246.

²⁶¹ São as características como a autonomia, a complexidade, a conectividade, a opacidade dos processos de decisão e a imprevisibilidade dos sistemas de inteligência artificial que dificultam a identificação dos responsáveis bem como a verificação do respetivo nexo causal, essencial para a obrigação de indemnizar. Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.32 e seguintes.

²⁶² Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 315.

seja, mesmo que aquele demonstre não ter agido com qualquer tipo de culpa e mesmo que até prove que teve a conduta mais diligente possível, irá responder pelos danos que originem da sua conduta pelo simples facto de esta consubstanciar um risco para o bem-estar geral da sociedade. Apesar da sua diligência, justifica-se a sua responsabilização pois a,quele que retira os benefícios e colhe os correspondentes proveitos, deve também de suportar os inerentes incómodos, tal como dita o princípio “*ubi commoda ibi incommoda*”²⁶³.

Ao se atribuir uma maior eficácia aos modelos da responsabilidade objetiva na procura de uma solução para esta incógnita jurídica, não se quer dizer que a responsabilidade civil subjetiva, assente na culpa, seja completamente inaplicável e desnecessária, muito pelo contrário, vai estar presente numa ótica de complementaridade. (“*Fault Liability (...) as well as Strick Liability for risks and for defective products, should continue to coexist*”²⁶⁴). Importante compreender que, mesmo que em minoria, “*em muitos casos se pode desvelar a culpa – por exemplo, porque o utilizador/operador não cumpriu determinados deveres de cuidado no sentido da atualização do software ou no sentido de impedir o acesso de terceiros – hackers – ao sistema*”²⁶⁵, portanto o legislador poderá ter o cuidado de prever um conjunto de deveres aos quais os programadores, operadores e outros colaboradores na cadeia de criação, desenvolvimento e manutenção dos sistemas de IA estejam adstritos, para que, na eventualidade de incumprimento destes mesmos deveres, se facilite o juízo de culpa e a sua prova²⁶⁶ (estratégia que parece ter sido adotada pela União, tal como será possível observar posteriormente).

Posto isto, não nos podemos esquecer que a responsabilidade objetiva é meramente excecional, tal como o n.º 2 do artigo 483.º o refere, “Só existe obrigação de indemnizar independentemente da culpa, nos casos especificados na lei”, o que significa que este princípio da tipicidade (*numerus clausus*), limita a nossa potencial solução (“*responsabilizar um ente (seja humano ou humanoide) independentemente de culpa implica a existência de uma previsão expressa do legislador nesse sentido*”²⁶⁷) que ao momento não existe. Cabe assim ao legislador intervir no sentido de alargar o âmbito dos regimes já existentes através de alterações legislativas que não só acautelem os perigos atuais, como também estabeleçam

²⁶³ Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, pp.246.

²⁶⁴ Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.36.

²⁶⁵ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 315.

²⁶⁶ *Ibidem*, pp. 316.

²⁶⁷ Ana Rita Maia, *Ob. cit.*, pp.31 e 32.

bases fortes para se conseguir acompanhar as rápidas evoluções e modificações do mundo da inteligência artificial, sob pena de ficarmos “*desamparados pela inexistência de uma norma que solucione o problema concreto.*”²⁶⁸. Vejamos algumas das possibilidades.

3.3.1. Presunção Legal do Artigo 493.º CC

Antes mesmo de envergarmos pelos regimes de responsabilidade objetiva e, em concordância com o dito anteriormente, é possível encontrar-se uma hipótese de elevado potencial para responder à nossa questão, presente no artigo 493.º do CC. Nesta norma constam duas presunções legais que podem suscitar algumas dúvidas de aplicação a sistemas de IAs.

Em primeiro lugar, o n. 1 deste preceito legal estipula uma “*responsabilidade por culpa presumida*”, onde quem “*tiver em seu poder coisa móvel ou imóvel, com o dever de a vigiar, e bem assim quem tiver assumido o encargo da vigilância de quaisquer animais, responde pelos danos que a coisa ou os animais causarem, salvo se provar que nenhuma culpa houve da sua parte ou que os danos se teriam igualmente produzido ainda que não houvesse culpa sua*”²⁶⁹. Podemos configurar este regime como uma opção pela “*ausência da personificação e da atribuição das inerentes capacidades jurídicas*”²⁷⁰, ou seja, é a ausência de personalidade jurídica dos sistemas de IA, que nos permite equacionar como solução este regime dirigido a coisas móveis/imoveis e animais²⁷¹.

A questão principal que aqui se pode colocar está associada à existência de um eventual dever de vigilância sobre um sistema de IA, mas também associada à “*determinação da sua fonte, a sua extensão e mesmo a sua exequibilidade*”, “*sendo um robot (...) autónomo, nem sempre será fácil conceber uma forma ou obrigação de o vigiar*”. Mais uma vez, devido às características inatas destes entes tecnológicos, deparamo-nos com um obstáculo: a incompatibilidade da vigilância com a sua autonomia. Quanto a esta questão,

²⁶⁸ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 285.

²⁶⁹ Luís Menezes Leitão, *Ob. cit.*, p.326 e 327; Ana Elisabete Ferreira, “Responsabilidade Civil Extracontratual por danos causados por robôs autónomos: breves reflexões”, in *Revista Portuguesa do Dano Corporal*, ano XXV, n.º 27, dezembro, 2016, p.59.

²⁷⁰ Henrique Sousa Antunes, *Ob. cit.*, p.147.

²⁷¹ Temos efetivamente duas realidades plasmadas no n.º 1 do artigo 493.º: coisa móvel ou imóvel, e animais. Quanto a esta dupla conjectura, podemos e devemos de configurar sistemas de IA como coisas móveis/imóveis, e não como animais, já que estes algoritmos inteligentes se encontram com maior proximidade da inteligência humana do que os próprios animais, provando ser assim duas realidades bem distintas. Do ponto de vista da personificação jurídica “mais facilmente a IA deverá encontrar-se dela revestida do que um animal” (José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, p.98)

Ana Rita Maia apresenta uma posição onde defende que “*a aplicação deste regime aos entes dotados de IA pode proceder quando existe margem para a intervenção humana*”, assim, deve de se promover que “*em nome da segurança, todos os entes dotados de IA possuam um comando que permita desligar no imediato todo os sistemas em caso de perigo ou de risco iminente, permitindo a atuação do obrigado à vigilância*”²⁷².

Contudo, parece-nos difícil esta verificação do dever de vigilância ou monitorização, principalmente em sistemas de IAs dotados de um elevado nível de autonomia, ainda que se promova uma espécie de interruptor *ON/OFF*. Quanto maior for o nível de autonomia de um sistema, maioritariamente, em sistemas munidos de algoritmos de aprendizagem automática ou profunda, menor será a margem de atuação do homem, dificultando assim a legitimidade deste dever de vigilância. Este preceito “*definitivamente não foi pensado para robôs autônomos*”²⁷³. De acordo com a jurisprudência, podemos encontrar o fundamento deste dever de vigilância num “*dever geral de prevenção do perigo ou nos deveres de segurança do tráfego*”. Assim, o dever de vigilância que recai sobre o sujeito é gerado “*pela especial perigosidade da coisa*”²⁷⁴, o que nos leva para a averiguação do perigo que representa um sistema de IA, que como veremos de seguida, não é alvo de uma resposta simples.

Por sua vez, de acordo com o n.º 2 do artigo 493.º, o legislador estipulou que, “*quem causar danos a outrem no exercício de uma atividade, perigosa por natureza ou pela natureza dos meios utilizados, é obrigado a repará-los, exceto se mostrar que empregou todas as providências exigidas pelas circunstâncias com o fim de os prevenir*”. Mais uma vez, constatamos a inversão do ónus da prova a favor do lesado, recaindo sobre o lesante, quem exerceu a atividade perigosa, uma presunção que, sendo ilidível, só poderá ser afastada caso aquele mostre que empregou todas as providências exigidas e adequadas com o fim de prevenir os danos. É do entendimento da doutrina, a existência de uma certa proximidade deste preceito com o regime da responsabilidade objetiva, ficando bem claro que a prova da relevância negativa da causa virtual não funciona como uma forma de eximir a responsabilidade, sendo antes exigido a demonstração de que o agente “empregou todas as

²⁷² Ana Rita Maia, *Ob. cit.*, pp.37 e 38.

²⁷³ Ana Elisabete Ferreira, *Ob. cit.*, pp.62.

²⁷⁴ Ac. do Tribunal de Justiça de 30/09/2014, Processo n.º 368/04.0TCSNT.L1.S1, relator Maria Clara Sottomayor.

providências exigidas pelas circunstâncias com o fim de prevenir os danos”²⁷⁵. Esta é uma “*norma particularmente flexível*”²⁷⁶.

A questão que aqui se pode colocar é quanto à definição de “atividade perigosa por sua própria natureza ou pela natureza dos meios utilizados”, bem como a correspondente aplicação aos sistemas de inteligência artificial.

O conceito de “atividade perigosa” que o legislador nos apresenta, é um conceito aberto e indeterminado, relacionando apenas a perigosidade com a natureza da atividade ou com os meios utilizados²⁷⁷. Assim, recorrendo à jurisprudência, concluímos que uma atividade será perigosa, segundo o Tribunal da Relação de Lisboa²⁷⁸, se “*em razão da sua própria natureza ou pela natureza dos meios utilizados, a ela andar ligada a ameaça de danos a terceiros, a forte probabilidade de lesões ou a potencialidade de provocação de danos (...) pressupõe, assim, uma probabilidade de aquela concreta atividade desenvolvida pelo lesante, causar um dano a terceiro (...) a concreta atividade desenvolvida pelo lesante acarrete um perigo que exceda, que vá para além do que é normal noutras atividades, que seja expectável que dela possam resultar danos que, em termos de normalidade, não ocorreriam noutra atividade*”.

No que aos sistemas de inteligência artificial diz respeito e quanto ao seu enquadramento como uma atividade perigosa por sua própria natureza ou pela natureza dos meios utilizados, a resposta não se afigura simples. Segundo alguns autores, não “*parece seguro concluir que a utilização de robots será necessariamente uma atividade perigosa*”, até porque estes sistemas já são “*genericamente melhores do que os seres humanos*”²⁷⁹ na execução de variadas tarefas, apresentando resultados mais eficazes e céleres, com “*menor margem de erro*”, logo, com menor probabilidade de perigo²⁸⁰. Outros, como Henrique Antunes, já entendem que a perigosidade também deve de aferir o grau de envolvimento da atividade com os bens pessoais que serve, deste modo “*quanto maior a proximidade dessa conduta (...) a bens existenciais, maior a probabilidade de um dano grave*”²⁸¹. Assim, parece-nos adequado considerar, tal como Ana Rita Maia sugere, que a aplicação deste

²⁷⁵ Luís Menezes Leitão, *Ob. cit.*, p.2.

²⁷⁶ Nuno Sousa e Silva, *Ob. cit.*, p.521; também Ana Rita Maia, *Ob. cit.*, p.34.

²⁷⁷ Ana Elisabete Ferreira, *Ob. cit.*, pp.60.

²⁷⁸ Ac. do Tribunal da Relação de Lisboa de 22/06/2021, relator Jorge Capacete, Processo n.º 1694/18.6T8PDL.L1-7.

²⁷⁹ Nuno Sousa e Silva, “Direito e Robótica: uma primeira aproximação”; e Nuno Sousa e Silva “Inteligência Artificial, robots e responsabilidade civil: O que é diferente?”, in *Revista de Direito Civil*, ano IV, n.4, Almedina, 2019, p.702.

²⁸⁰ Ana Rita Maia, *Ob. cit.*, p.35.

²⁸¹ Henrique Sousa Antunes, *Ob. cit.*, p.146.

regime aos sistemas de inteligência artificial, é “*aceitável*”, contudo, terá de haver uma certa ponderação quanto à categorização de IAs como atividades perigosas (terá de ser um sistema de IA que possua uma “*especial aptidão produtora de danos*”²⁸²), sob pena de, por excesso, “*paralisar o avanço tecnológico*”²⁸³. “*Pense-se o quão estranho seria considerar a utilização de um smartphone uma atividade perigosa*”²⁸⁴.

As presunções que aqui residem no artigo 493.º, aproximam-nos de uma potencial solução ao nosso problema, contudo, é precisamente no seu ponto forte que encontramos mais alguns entraves à sua aplicação. Esta inversão do ónus da prova, permite-nos escapar à enorme dificuldade que recai sobre o lesado (provar a culpa do autor da lesão de acordo com o artigo 487.º/n.º 1), contudo, também é problemática, pois, como se tratam de presunções ilidíveis (artigo 350.º/n.º 2 do CC), deixam aberta a possibilidade de o lesante eximir-se da responsabilidade caso prove que não houve culpa nenhuma da sua parte ou que os danos se teriam igualmente produzido ainda que não houvesse culpa sua, para as situações do n.º 1, ou prove que empregou todas as providências exigidas pelas circunstâncias com o fim de prevenir os danos, para as situações do n.º 2. Assim, parece haver maior probabilidade de deixar-se o lesado desamparado com o afastamento da presunção de culpa, aspeto que os regimes de responsabilidade objetiva parecem tutelar melhor, ao prescindirem desse elemento. “*Deverá sempre a norma permitir a indemnização do lesado mesmo quando se mostrem empregues todas as medidas aversivas ao dano*”, pois, mesmo com um elevado grau de diligência, a imprevisibilidade e autonomia, características comuns de um sistema de IA, podem sempre culminar em danos²⁸⁵, o que significaria que teríamos uma grande válvula de escape de responsabilidade nestes casos.

Importante não esquecer que, esta presunção de culpa do artigo 493.º, não dispensa a prova do nexo de causalidade, pressuposto este que ainda terá de ser demonstrado, persistindo, dessa forma, a dificuldade sentida nesse âmbito.

²⁸² No Ac. do Tribunal de Justiça de 17(05(2017, Processo n.º 150/11.1TBOAZ.P1.S1, relator António Piçarra, estipulou-se que “*é atividade perigosa, para o efeito, aquela que possui uma especial aptidão produtora de danos, um perigo especial, uma maior suscetibilidade ou aptidão para provocar lesões de gravidade e mais frequentes, e que essa perigosidade deve ser aferida a priori e em abstrato e não em função dos resultados danosos, em caso de acidente, muito embora a magnitude destes possa evidenciar o grau de perigosidade da atividade ou risco dessa atividade*”

²⁸³ Ana Rita Maia, *Ob. cit.*, p.35.

²⁸⁴ *Ibidem*.

²⁸⁵ *Ibidem*, p.36.

3.3.2. Artigo 500.º CC

Uma outra solução para esta problemática dos sistemas de IA poderá encontrar-se na responsabilidade objetiva que se verifica nas relações de comitente-comissário. É assim defendida uma interpretação extensiva deste regime de responsabilidade objetiva do artigo 500.º como uma potencial solução para a questão dos danos causados por sistemas de inteligência artificial ²⁸⁶.

Partindo do anteriormente mencionado ²⁸⁷, a responsabilidade objetiva do comitente está dependente da verificação cumulativa de três requisitos: 1) a existência de uma relação de comissão; 2) o comissário tem de praticar os factos danosos no exercício da função que lhe foi confiada pelo comitente; 3) e por fim, tem de recair igualmente sobre o comissário a obrigação de indemnizar. Só com a verificação cumulativa de todos estes requisitos é que o comitente responderá objetivamente, podendo depois gozar de um direito de regresso para com o comissário.

Adaptando esta possibilidade aos sistemas de inteligência artificial temos de percorrer esses mesmos pressupostos. Quanto à existência de uma relação de comissão, poderá o comissário ser um ente tecnológico? Que “*vínculo se há de conjeturar existir com o comitente?*” ²⁸⁸. Não parece errado dizer que este requisito se pode encontrar verificado, na medida em que estes sistemas podem estar a ser utilizados, a prosseguir ou executar certas tarefas e atividades por conta e no interesse de outrem, passando assim este primeiro ponto de controlo.

Face ao segundo requisito, ou seja, a indispensabilidade de o comissário provocar os danos no exercício da função que lhe foi confiada ²⁸⁹, também é imaginável a sua verificação, embora possa depender muito do estado tecnológico da ciência. Ao dia de hoje, apesar dos elevados níveis de automatismo e grandes avanços tecnológicos, ainda nos encontramos muito limitados com os sistemas de inteligência artificial. Talvez um dia haja uma maior deliberação das máquinas com recurso ao *machine learning* e *deep learning*, mas parece que até ao momento o seu funcionamento é mais linear, no sentido de seguirem um “*determinado itinerário*” ou codificação, ao invés de atuarem de maneira arbitrária e contrária às instruções

²⁸⁶ Tal como o artigo 11.º do CC nos leva a querer “*As normas excepcionais não comportam aplicação analógica, mas admitem interpretação extensiva.*”.

²⁸⁷ Questão discutida e desenvolvida no ponto 3.2.2.1.

²⁸⁸ José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, p.98 e 99.

²⁸⁹ Cf. artigo 500.º/n.º 2, 1.º parte.

inicialmente dadas. Parece, pois, precoce, pensar que uma IA, como comissário, se desviasse das funções que lhe fossem confiadas. “*O Prolema, porém, há de pôr-se rapidamente.*”²⁹⁰.

Por fim, no último requisito temos o maior obstáculo a uma possível interpretação extensiva deste regime, sendo também por esta razão que se afasta de todo esta possibilidade. De acordo com a parte final do n.º 2 do artigo 500.º, para que o comitente responda objetivamente pelos danos praticados pelo comissário, terá de recair sobre este último uma obrigação de indemnizar, ou seja, o comissário tem de ser ele próprio responsável, o que significaria que ele responderia, exclusivamente, pelo dano causado a terceiro, não fosse a relação de comissão. Torna-se, portanto, “*fundamental (...) que os requisitos de alguma espécie de responsabilidade civil estejam preenchidos contra o comissário*”²⁹¹. Ainda que seja possível estarmos a falar de uma modalidade de responsabilidade pelo risco ou por factos lícitos, em regra estará sempre em causa a típica responsabilidade subjetiva que em princípio vai pressupor a culpa, o que nos leva de volta ao problema sentido anteriormente²⁹². Tal como já referido num momento pretérito, estes sistemas “*não sendo pessoas*”, não tendo personalidade jurídica, “*não são centros autónomos de imputação*”, logo não faz sentido “*serem alvo de juízos ético-jurídicos de censura por uma conduta deficiente*”, levando à não verificação deste último pressuposto.

Por estas mesmas razões, somos forçados a considerar esta solução como “*imprestável*” e insuficiente para solucionar o dilema da responsabilidade pelos danos causados por sistemas de inteligência artificial, pelo simples facto de não estarem verificados todos os requisitos legais necessários à sua aplicação^{293 294}.

3.3.3. Responsabilidade do produtor por produtos defeituosos

Uma outra possível solução para a problemática aqui em causa, seria a aplicação do regime da responsabilidade objetiva do produtor pelos danos provocados pelos produtos defeituosos que colocou em circulação, presente no decreto-lei n.º 383/89 de 6 de novembro,

²⁹⁰ José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, pp.101 e 102.

²⁹¹ *Ibidem*, p.102/103.

²⁹² *Ibidem*.

²⁹³ Manuel Felício, “Responsabilidade Civil por acidente de viação causado por veículo automatizado”, em *Revista de Direito da Responsabilidade*, ano 1, 2019, pp. 514 e 515.

²⁹⁴ “*parece improceder a este nível, por falta a possibilidade de se imputar o evento lesivo ao comportamento do robô, exceto se o quisermos personificar*”, Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, p.287.

que transpôs para a ordem jurídica interna a Diretiva n.º 85/374/CEE, do Conselho, de 25 de julho de 1985.

Tal como já tivemos a oportunidade de expor anteriormente, para que este regime de responsabilidade objetiva do produtor se possa aplicar é necessário, para além dos requisitos normais da responsabilidade objetiva, a verificação de requisitos especiais delimitados pela diretiva, a saber os conceitos de produtor, produto e defeito (sendo ainda relevante o momento de entrada do produto em circulação). No que à Inteligência Artificial diz respeito, temos de saber se as suas especificidades se subsumem nos âmbitos desta legislação.

Em primeiro lugar, mencionamos que o conceito de produtor que aqui encontramos é um tanto amplo, englobando três possíveis sentidos: o produtor real, produtor aparente e produtor presumido ²⁹⁵. Assim, considerando o supramencionado juntamente com o contexto da nossa temática, não parece surgirem dúvidas quanto à verificação deste pressuposto. De acordo com estas definições, teremos como produtor de IA “*entre outros, o engenheiro robótico, o programador ou produtor de software, o produtor de hardware ou aquele que apenas tenha recebido o robô de um fornecedor com o objetivo de o vender, alugar ou qualquer outra forma de distribuição.*” ²⁹⁶.

O passo que se segue é saber se um sistema de inteligência artificial se enquadra no produto a que se refere a diretiva, “*qualquer coisa móvel, ainda que incorporada noutra coisa móvel ou imóvel.*” ²⁹⁷. Num primeiro momento e, de acordo com o n.º 1 do artigo 202.º do código civil, “*diz-se coisa tudo aquilo que pode ser objeto de relações jurídicas.*” ²⁹⁸, já quanto à sua qualificação como “coisa móvel” o artigo 205.º/n.º 1 utiliza um critério de exclusão, ou seja, são móveis todas as coisas não compreendidas no artigo que define e elenca as coisas imóveis ²⁹⁹. Quanto à verificação deste quesito, se olharmos para uma inteligência artificial como o conjunto de um *software* embutido num *hardware* (num suporte físico), podemos sim considerá-lo uma coisa móvel complexa, contudo, as dúvidas crescem quando falamos de um sistema de inteligência artificial só como um algoritmo ou

²⁹⁵ Cf. art. 2.º do Decreto-Lei n.º 383/89, de 6 de novembro.

²⁹⁶ Juliana Campos, “A Responsabilidade Civil do Produtor pelos danos causados por robôs inteligentes à luz do regime do Decreto-Lei n.º 383/89, de 6 de novembro”, em *Revista de Direito da Responsabilidade*, ano 1, 2019, p.708.

²⁹⁷ Cf. art. 3.º do Decreto-Lei n.º 383/89, de 6 de novembro.

²⁹⁸ Esta é uma definição que levantou alguma discussão por ser um tanto ampla e imprecisa visto que nem tudo o que pode ser objeto de relações jurídicas é uma coisa, como por exemplo as pessoas, as prestações, entre outras. Mota Pinto apresenta-nos uma definição jurídica mais completa ao considerar coisas “*bens (ou entes) de caráter estático, desprovidos de personalidade e não integradores do conceito necessário desta, suscetíveis de constituírem objeto de relações jurídicas*”. Carlos Alberto da Mota Pinto, “*Teoria Geral do Direito Civil*”, Gestlegal, 2020, p.341.

²⁹⁹ Cf. art 204.º.

software, devido à sua imaterialidade³⁰⁰. A este propósito Mafalda Miranda Barbosa refere que para este efeito, a sua qualificação como incorpórea não é relevante, permitindo assim considerar-se um *software* como “o produto” a que a diretiva faz referência³⁰¹. Também Calvão da Silva refere que a definição de produto que consta na diretiva “*abrange os suportes materiais em que a obra intelectual se materializa, fixa e comunica, pois, são coisas móveis corpóreas, embora inconfundíveis com a obra intelectual em si – bem material*”, deste modo considera que “*programas estandardizados de computador (computer software) são, portanto, produtos no sentido do DL n.º 383/89.*”³⁰², justificando esta posição com a necessidade de proteção do consumidor face aos elevados riscos provenientes da comercialização dos *softwares*³⁰³.

Por último, o produtor só se vai responsabilizar caso o dano sofrido tenha sido provocado por um produto defeituoso, ou seja, segundo o artigo 4.º, um produto que “*não oferece a segurança com que legitimamente se pode contar*”. Portanto, não se exige que o produto seja totalmente seguro, muito pelo contrário, exige-se que o consumidor, atendendo às expectativas do público em geral e especificidades do caso em concreto, possa contar com um certo nível de segurança³⁰⁴, ou seja, haverá defeito se, tendo em atenção alguns aspetos como a sua apresentação, utilização que dele razoavelmente possa ser feita e momento da sua entrada em circulação, este apresente um risco inesperado, não sendo algo com que o consumidor possa legitimamente contar³⁰⁵. Posto isto, torna-se pertinente questionar qual será a segurança com a qual o consumidor pode legitimamente contar de um sistema de inteligência artificial já que este pode ter uma capacidade evolutiva elevada. Juliana Campos enumera várias situações em que esta definição de segurança se torna “*imprecisa*”, como o facto de os *softwares* seres suscetíveis de receberem atualizações após o momento da sua entrada em circulação, que podem claramente sanar o vício preteritamente verificado ou

³⁰⁰ Juliana Campos, *Ob. cit.*, p. 709; também Vera Lúcia Coelho, *Ob. cit.*, p.16;

³⁰¹ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, p.287.

³⁰² João Calvão da Silva, *Ob. cit.*, p.613; Neste sentido também Nuno Devesa Neto, “Responsabilidade Civil pela utilização de robots-advisors: A insuficiência do atual regime de responsabilidade e a necessidade de uma nova hipótese de responsabilidade objetiva”, em *Revista de Direito da Responsabilidade*, ano 2, 2020, p.922.

³⁰³ João Calvão da Silva, “*Compra e Venda de Produtos Defeituosos (Conformidade e Segurança)*”, Almedina, Coimbra, 2008, p.185.

³⁰⁴ “Defectiveness is assessed on the basis of the safety expectations of an average consumer”, Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.28.

³⁰⁵ Se por outro lado e, atendendo às especificidades mencionadas no n.º 1 do artigo 4.º, o risco for conhecido e “*tido por normal das conceções do tráfico do respetivo setor de consumo*”, então é considerado uma segurança com que legitimamente se pode contar, logo para este efeito não será um produto defeituoso. João Calvão da Silva, “*Responsabilidade Civil do Produtor*”, p.636/637.

alterar por completo o sistema em causa ³⁰⁶. Mais, qual a segurança com que se pode legitimamente contar perante uma IA que interaja com o seu meio ambiente, aprenda por si própria e, de acordo com a análise dos dados que recolheu, atinja uma conclusão totalmente autónoma e imprevisível. “*Note-se que, o estado da ciência e da técnica, aferido ao tempo da sua emissão no mercado, ainda não permite saber tudo aquilo que um robô vai aprender e, conseqüentemente, se este oferecerá um bom nível de segurança*” ³⁰⁷.

No entanto, ainda encontramos mais algumas especificidades. Tal como mencionado anteriormente ³⁰⁸, podemos qualificar este defeito como um defeito de conceção (erros ou deficiências na fase de idealização ou *design*), um defeito de fabrico (problemas na fase de laboração, produção, concretização do idealizado) ou um defeito de informação (falta, insuficiência ou inadequação de informações ou instruções de uso, por parte do produtor). Partindo destas considerações, podemos deparar com dois potenciais cenários.

O primeiro envolve um sistema de inteligência artificial defeituoso, contendo quer defeitos de conceção, quer de fabrico (“*falha no planeamento ou por um erro na fase de laboração, se preterem regras de segurança (...), má programação, inexistência de botões de segurança, não colocação de sensores, não implementação de mecanismos de redução da vulnerabilidade do software em relação à interferência de terceiros, entre outros exemplos*” ³⁰⁹) ou até mesmo defeito de informação (“*falta de instruções sobre o produto, a falta de advertência sobre o perigo, a falta de esclarecimento acerca das medidas de cuidado a ter em conta*” ³¹⁰). Assim, perante uma inteligência artificial defeituosa e, provando-se os “*danos efetivos e a relação de causalidade entre os danos e o defeito*” ³¹¹, aplicar-se-á este regime de responsabilidade objetiva do produtor ³¹².

O segundo cenário possível, comporta uma situação de maior dúvida quanto à aplicação deste regime. Trata-se das situações onde os danos originam não devido a um

³⁰⁶ Juliana Campos, *Ob. cit.*, p. 711.

³⁰⁷ *Ibidem*.

³⁰⁸ No ponto 3.2.2.5.

³⁰⁹ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 288.

³¹⁰ *Ibidem*.

³¹¹ De acordo com os considerandos AE e AH da Resolução do Parlamento Europeu de 2017 que referem respetivamente que “*a responsabilidade pelo produto – segundo a qual o fabricante de um produto é responsável por uma anomalia (...) são aplicáveis aos danos causados pelos robôs ou pela IA;*”, e “*no que respeita à responsabilidade extracontratual, a Diretiva 85/374/CEE, apenas pode abranger os danos provocados por defeitos de fabrico de um robô*”. Resolução do Parlamento Europeu de 16 de fevereiro de 2017 (2015/2103(INL)).

³¹² Juliana Campos, *Ob. cit.*, p. 715 e 716. Do mesmo modo afirma Ana Rita Maia “apenas se poderá considerar defeito, e, por conseguinte, aplicar-se o regime da responsabilidade objetiva do produtor, quando se consiga comprovar que os danos causados se fundam em defeitos de conceção, de informação carreada para o ente dotado de IA, e no seu fabrico.” Ana Rita Maia, *Ob. cit.*, p.18 e 19.

defeito do sistema de inteligência artificial, mas ocorrem como um efeito da sua atuação autónoma, ou seja, resultam do seu normal funcionamento, “*reconduz-se a uma ideia de “não é defeito, é feito”*.”³¹³. Nestes casos, já se levantam entraves à aplicação da Diretiva pelos danos surgirem devido a uma característica intrínseca da IA, a autonomia, ao invés de “*ser uma marca de defeituosidade*”^{314 315}. Podemos assim concluir que, estes sistemas de IA, munidos de uma capacidade autónoma elevada e com avançados algoritmos de *machine learning e deep learning*, se enquadram dentro de uma “nova geração de robôs”, situação para a qual o “atual” quadro jurídico da Diretiva 85/374/CEE, já não seria suficiente para abranger os danos provocados por eles^{316 317}. Partilhamos assim, esta ótica negativa de aplicação da diretiva para uma geração de sistemas de IA novos, mais autónomos e independentes (baseados em algoritmos de autoaprendizagem), pelo menos enquanto não se observam alterações relevantes aos seus conceitos (como por exemplo, a modificação do conceito de defeito para fazer referência à noção de autonomia³¹⁸)³¹⁹.

Uma outra situação que merece atenção está ligada aos riscos de desenvolvimento, uma das causas de exclusão da responsabilidade previstas no artigo 5.º do decreto-lei. Apesar da verificação de todos os pressupostos, o produtor pode eximir-se da responsabilidade objetiva caso prove que “o estado dos conhecimentos científicos e técnicos, no momento em que pôs o produto em circulação, não permitia detetar a existência do defeito”. Ora, este “escudo de imunidade”³²⁰ poderá ser problemático, já que esta área das inovações tecnológicas sofre com umas das suas grandes virtudes, a rapidez (é positiva a velocidade com que se atingem grandes avanços e descobertas em prol da ciência, da medicina, do bem

³¹³ Juliana Campos, *Ob. cit.*, p. 712.

³¹⁴ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 288.

³¹⁵ “*Sophisticated AI autonomous systems with self-learning capabilities also raise the question of whether unpredictable deviations in the decision-making path can be treated as defects*”, Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.28.

³¹⁶ *Ibidem*, considerando AI.

³¹⁷ Importante ter em conta que, como se está a lidar com entes tecnológicos, que podem alterar não só com a interação com o meio, mas também devido a novas instruções, *updates* que possam vir a receber, é possível verificar-se *a posteriori* qualquer defeito novo, contudo, o produtor pode eximir-se de qualquer responsabilidade se provar que “tendo em conta as circunstâncias, no momento da entrada do produto em circulação, inexistia qualquer defeito, de acordo com a alínea b) do artigo 5.º. Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 289; contudo, este aspeto ainda levanta algumas críticas que serão mencionadas mais à frente no texto.

³¹⁸ Juliana Campos, *Ob. cit.*, p. 719.

³¹⁹ “*Precisamente devido à “free will” (maior ou menor) de que goza a máquina dotada de IA, será de excluir liminarmente o recurso ao instituto da responsabilidade (objetiva) do produtor (Decreto-Lei n.º 383/89, de 6 de novembro) (...). A menos que a respetiva manifestação se conceba como um defeito.*” José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, pp.90 e 91.

³²⁰ Juliana Campos, *Ob. cit.*, p. 716.

social, do bem laboral, bem como de outros aspetos de lazer, mas também pode ser negativa pela grande imprevisibilidade que acarreta, pois esse desconhecimento pode levar à perda de controlo de um conjunto de circunstâncias importantes – por exemplo, questões de responsabilidade ou privacidade – devido a um abuso e excesso de tecnologia). Com isto quer-se dizer que o estado da arte, ciência e da técnica hoje, será sempre inferior ao de amanhã e, como “*não se conseguem testar todas as combinações possíveis que um robô pode aprender*”, estes sistemas nunca estarão livres de riscos ³²¹. Sendo assim, quando estiverem em causa sistemas de IA e eventuais danos causados por estes, podemos nos deparar muitas vezes com defeitos ou imprecisões que o estado científico e técnico, no momento da entrada em circulação do produto, não permitia conhecer, o que significa que o produtor pode facilmente eximir-se da sua responsabilidade ao abrigo desta alínea e) (“*esta cláusula terá uma aplicação exponencial quando o produtor conceber os robôs dotados de IA autónomos*” ³²²).

É devido a estes contornos que a permanência dos riscos de desenvolvimento se torna controversa para alguns autores. Há quem defenda que este preceito dos riscos de desenvolvimento deve de ser suprimido pois “*serão férteis os casos em que se legitimará a sua aplicação*” ³²³. No mesmo sentido, Sónia Moreira partilha de uma posição que consta do relatório do Expert Group on Liability and New Technologies e que refere que “*A development risk defence should not apply.*” ³²⁴, concordando que “*atendendo à necessidade de distribuição justa dos riscos e dos benefícios, não deve permitir-se ao produtor escusar-se da sua responsabilidade invocando o risco de desenvolvimento. É que, neste caso, é previsível que desenvolvimentos imprevisíveis possam ocorrer.*” Já por outro lado, Ana Rita Maia, discorda da abolição dos riscos de desenvolvimento pois, “*aceitar sem qualquer reserva este argumento pode prejudicar os avanços tecnológicos, pois o produtor estaria ad eternum sujeito à imputação de uma responsabilidade (...) da qual eximir-se seria custoso*” ³²⁵. Reconhecendo que ambas as posições têm os seus lógicos argumentos, parece ser mais adequado uma visão baseada numa adaptação do conceito de riscos de desenvolvimento, para que melhor se aborde esta questão dos sistemas de IA, ou, como

³²¹ *Ibidem*, p.719.

³²² *Ibidem*, p.716.

³²³ *Ibidem*, p.719.

³²⁴ Sónia Moreira, “Considerações sobre Inteligência Artificial e Responsabilidade Civil: O caso dos Veículos Autónomos”, em Maria Miguel Carvalho, *E-Tec Yearbook – Artificial Intelligence & Robots*, School of Law of the University of Minho, dezembro, 2020, p.86. E Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.43.

³²⁵ Ana Rita Maia, *Ob. cit.*, p.21.

Mafalda Miranda Barbosa defende, a solução talvez passe não por uma alteração legislativa, mas sim por uma “*adequada mobilização do que é o risco de desenvolvimento*”, pois mesmo que o sistema de IA seja autónomo e imprevisível, “*é possível prever que uma lesão viesse a ocorrer, exatamente porque a aprendizagem pela interação do meio do algoritmo seria imprevisível.*”³²⁶.

Posto isto, quanto à adequação deste regime como uma potencial solução para o problema colocado pelos sistemas de inteligência artificial, é importante referir que a União Europeia, já em 2017/2018, estava a proceder a estudos e avaliações que permitissem dar início a uma tentativa de adaptação dos regimes legais de responsabilidade existentes às novas tecnologias. Resultou destes esforços, a Avaliação da Diretiva 85/374/CEE, do Conselho, de 25 de julho de 1985³²⁷, que teve como principal objetivo avaliar o funcionamento, a *performance* e a eficácia desta legislação de um modo geral, durante o período de 2000 a 2016. Contudo, também pretendiam, de certa forma, tirar conclusões quanto à eficácia de uma possível aplicação e adaptação do regime face às tecnologias digitais emergentes³²⁸. Por um lado, de um modo geral, e com base nos resultados apresentados, a grande maioria dos produtores e consumidores questionados afirmaram satisfação e confirmaram a adequação da Diretiva para acautelar as suas necessidades quando estivessem em causa desenvolvimentos tecnológicos inovadores. Por outro lado, surgiram algumas dúvidas em relação a uma categoria diferente de produtos que identificaram como potenciais problemas ou incertezas quanto à aplicação da diretiva, entre eles o nosso caso de sistemas de inteligência artificial (“*products where software and applications from different sources can be installed after purchase, products performing Automated task based on algorithms, data analytics, self-learning algorithm or products purchased as a bundle with related services*”), ou seja, estes produtos podem tornar-se

³²⁶ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 322.

³²⁷ SWD(2018) 157 final.

³²⁸ Avaliação da Diretiva 85/374/CEE do Conselho, de 25 de julho de 1985, relativa à aproximação das leis, regulamentos e disposições administrativas dos Estados-Membros relativas à responsabilidade por produtos defeituosos, (SWD(2018) 157 final.), Comissão Europeia, Bruxelas, 2018, p.3.

problemáticos na aplicação da diretiva devido ao seu elevado grau de complexidade e autonomia^{329 330 331}.

Torna-se assim necessário uma revisão e clarificação dos conceitos essenciais da diretiva para que esta se adeque melhor às situações englobadas pelo seu âmbito, mas para que também se consiga encontrar respostas para um número vasto de problemas relacionados com sistemas de inteligência artificial e outras novas tecnologias.

Deste modo o conceito de produtor, essencial para esta responsabilidade objetiva, pode tornar-se cada vez menos claro devido aos casos particulares das tecnologias digitais emergentes, uma dificuldade que poderia ser ultrapassada com um alargamento da definição a todos os potenciais intervenientes no processo de produção e manutenção³³².

Também o conceito de produto terá de ser alvo de algumas alterações principalmente para sanar quaisquer dúvidas que ainda possam existir quanto à consideração dos *softwares* e restantes sistemas de IAs como produtos, bem como quaisquer componentes digitais que ainda possam suscitar problemas³³³. Estas novas tecnologias digitais tornam a linha que separa produtos e serviços muito ténue³³⁴. Questiona-se também, se estariam abrangidos no conceito de “produto” outros conteúdos digitais que, apesar de funcionarem de maneira equivalente aos *softwares*, não executem tarefas por si próprios, ou se, a comercialização de um *software* como um serviço, “*SaaS (Software-as-a-Service)*”³³⁵, também permitiria a respetiva compensação pelos danos provocados, nos mesmo termos. Portanto, uma clarificação da definição permitir-nos-ia saber se os danos provocados por esse tipo de sistemas inteligentes também estariam abrangidos, já que do ponto de vista do consumidor

³²⁹ *Ibidem*, pp. 32 a 35.

³³⁰ Observação que já tinha sido referenciada nesta investigação e também pela Resolução do Parlamento Europeu de 16 de fevereiro de 2017 (2015/2103(INL)), quando no seu considerando AI, se pode ler que “*não obstante o âmbito de aplicação da Diretiva 85/374/CEE, o atual quadro jurídico não seria suficiente para abranger os danos provocados pela nova geração de robôs, na medida em que os robôs podem ser dotados de capacidades cognitivas e de aprendizagem que integram um certo grau de imprevisibilidade no seu comportamento, uma vez que aprendem de forma autónoma com a sua experiência própria variável e interagem com o seu ambiente de um modo único e imprevisível.*”

³³¹ Assim, o relatório concluiu que, seria necessário a realização de uma avaliação mais rigorosa e detalhada, com uma atenção especial dada aos desenvolvimentos tecnológicos, sendo muito provável a necessidade de alguma clarificação dos conceitos e adaptação de muitos outros aspetos à nova realidade verificada. É referido ainda que, devido à falta de dados empíricos, ou seja, devido à falta de incidentes práticos significativos que demonstrassem a ineficácia da diretiva, seria preciso obter-se mais provas para se determinar se a legislação ainda se adequaria ao contexto envolvente ou não.

³³² (SWD(2018) 157 final.), p.61.

³³³ *Ibidem*, p.61.

³³⁴ Torna-se “*difícil estabelecer a cisão clara entre produtos e serviços, na medida em que produtos e serviços interagem continuamente*”, Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 291.

³³⁵ Teresa Rodríguez Ballell, “The revision of the product liability directive: a key piece in the artificial intelligence liability puzzle”, *ERA Forum*, vol. 24, n.º 2, junho, 2023, pp. 253 e 254.

não haveria diferença na procura da sua justa compensação (“*from the perspective of the victim no distinction in the commercialisation model – provided as a standalone product or under a subscription agreement*”) ³³⁶.

Por sua vez, o próprio conceito de defeito também se apresenta relativamente desapropriado ao contexto dos sistemas de inteligência artificial. É “*incompatível com a ideia de autonomia*” ³³⁷, ou seja, este pressuposto apenas permite a aplicação da diretiva aos casos onde os danos provocados por IAs, origem do defeito, tal como ele é compreendido no diploma, e não como efeitos das suas características intrínsecas. Assim, para responder aos cenários onde os danos provocados devem-se à “*autonomia deliberativa*” ³³⁸ dos sistemas de inteligência artificial, seria preciso “*existir uma modificação do conceito de defeito (...) a qual deveria fazer referência à noção de autonomia*” ³³⁹. Embora a posição de Juliana Campos faça todo o sentido numa ótica de alargamento do âmbito de aplicação da diretiva ³⁴⁰, parece-nos uma tarefa que exige um elevado nível de meticulosidade para que não se obste ao bom desenvolvimento da tecnologia. “*Se pretendemos máquinas autocéfalas, não podemos depois entender que o exercício da autonomia se tenha como uma deficiência quando, porventura, se desenvolver em sentidos, à partida, imprevistos*” ³⁴¹.

Dos resultados dos estudos, outro aspeto que é criticável nesta conceção da diretiva prende-se com a limitação temporal que a mesma coloca à verificação do defeito, mais precisamente “o momento da entrada do produto em circulação” ³⁴². Ao contrário dos produtos tradicionais, os sistemas de IA ou outro tipo de produtos munidos de *software* inteligente estão dependentes de várias informações (“*data-dependent*”) quer seja pelas instruções dadas pelo seu produtor, personalizações realizadas pelo utilizador ou todo um outro tipo de informações que recolham da sua interação com o meio ³⁴³. Fora isto, normalmente estarão suscetíveis a atualizações (*updates*) que permitem um aperfeiçoamento do *software* a vários níveis. Ora, isto poderá ser problemático na medida em que, após um destes produtos entrar em circulação, estando originalmente desprovido de qualquer defeito,

³³⁶ *Ibidem*.

³³⁷ Juliana Campos, *Ob. cit.*, p. 718.

³³⁸ Ana Rita Maia, *Ob. cit.*, p.18.

³³⁹ Juliana Campos, *Ob. cit.*, p. 719.

³⁴⁰ Em concordância com “The solution has been to extend the list of factors to be considered in the assessment of the defectiveness of a product.”, Teresa Rodríguez Ballell, *Ob. cit.*, p. 255.

³⁴¹ José Alberto González, “Responsabilidade por danos e Inteligência Artificial (IA)”, p.91.

³⁴² É uma limitação temporal pois de acordo com a alínea b) do artigo 5.º, umas das possibilidades que o produtor tem de se eximir da responsabilidade é provar a inexistência do defeito no momento em que introduziu o produto em circulação. Sendo também este critério o utilizado por exemplo na exclusão que invoca os riscos de desenvolvimento.

³⁴³ Teresa Rodríguez Ballell, *Ob. cit.*, p. 257.

este possa vir demonstrar, *a posteriori*, qualquer deficiência, possivelmente proveniente desses *updates*. É por isto que o Expert Group recomendou que se considerasse o “*produtor responsável objetivamente pelos defeitos das tecnologias digitais emergentes, mesmo que tais defeitos surjam depois do produto ser colocado em circulação, desde que o produtor ainda estivesse em controlo das atualizações ou melhorias na tecnologia*”³⁴⁴. Para este efeito, teria de se repensar a relevância do momento da entrada em circulação do produto, pois esta “*perde sentido*”, já que com a realização de atualizações é como se “*continuamente o produtor estivesse a promover a entrada no mercado de produtos intangíveis, desmaterializados*”^{345 346}. Para além do mais, ainda acresce eventuais problemas quanto a quem deve de realizar a atualização do software, sendo que a resposta que se apresenta como mais sensata parece ser a perfilhada por Juliana Campos em que o “*produtor deve informar o utilizador ou proprietário das atualizações necessárias, pelo que se estes não a efetuarem e, por esse motivo, ocorrer um dano, haverá culpa do lesado podendo ser excluída a indemnização*”³⁴⁷.

Por fim, quanto aos danos indemnizáveis, por força do artigo 8.º do Decreto-lei n.º 383/89, só são indemnizáveis os danos resultantes de morte ou lesão pessoal e os danos em coisa diversa do produto defeituoso³⁴⁸, desde que esta seja normalmente utilizada para uso ou consumo privado e o lesado lhe tenha dado principalmente esse destino, ficando de fora os danos puramente patrimoniais (“*que podem avultar no contexto da utilização de entes dotados de inteligência artificial (...) uso de robot advisors ou no quadro da blockchain.*”³⁴⁹). Assim torna-se também necessário que se retoque o regime no sentido da inclusão dos

³⁴⁴ Tradução própria. Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019, p.42. No mesmo sentido Sónia Moreira, *Ob. cit.*, p. 86.

³⁴⁵ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 321.

³⁴⁶ Sendo os avanços e descobertas tecnológicas constantes, faz sentido o produtor manter um dever de vigilância ou sequela, “*não perdendo o rastro dos robots dotados de IA que produz*”, de acordo com o estipulado no Decreto-Lei n.º 69/2005, artigo 6.º/n.º 1, alínea b) “A tomar medidas apropriadas, em função das características do produto fornecido, à informação sobre o riscos que o produto possa apresentar e ao desencadeamento das ações que se revelem adequadas, incluindo a retirada do mercado, o aviso aos consumidores em termos adequados e eficazes ou a recolha do produto junto destes”. Ana Rita Maia, *Ob. cit.*, p.20; Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 289.

³⁴⁷ Juliana Campos, *Ob. cit.*, p. 712.

³⁴⁸ Os danos causados em coisas diversas do produto defeituoso também só são indemnizáveis na medida em que excedam o valor de (euro) 500. O que acaba por limitar bastante o número de pretensões. (SWD(2018) 157 final.), p.61.

³⁴⁹ Mafalda Miranda Barbosa, “O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução”, pp. 290. Também Nuno Devesa Neto, *Ob. cit.*, p. 924/925.

danos em coisas materiais, próprias ou de terceiros, não só quando utilizados para uso particular ³⁵⁰.

Em síntese, a conclusão a que chegamos é que esta Diretiva 85/374/CEE é vista com bons olhos como uma potencial solução para a problemática que os danos causados por sistemas de inteligência artificial levantam aos quadros jurídicos de responsabilidade civil (*“In the opinion of the NFT of the Expert Group, the principle of producer responsibility, adopted in relation to traditional products, should also apply to emerging digital technologies.”* ³⁵¹), contudo, o mesmo não é dizer que o diploma se mostra suficiente ou adequado à data das avaliações realizadas, principalmente quando nos deparamos com sistemas de IA desprovidos de qualquer defeito, onde este regime não terá qualquer aplicação. Com isto queremos dizer que, os princípios associados ao diploma são um bom ponto de partida, mas é necessária uma revisão de vários aspetos, para que se consiga não só responder às, ainda, dúvidas constantes (bem como novas dúvidas que são colocadas em prol da IA), como também para melhor se tutelar estes novos casos.

³⁵⁰ Ana Rita Maia, *Ob. cit.*, p.20.

³⁵¹ Expert Group on Liability and New Technologies, *“Liability for Artificial Intelligence and other emerging digital technologies”*, União Europeia, Luxemburgo, 2019, p.42.

4. Abordagem Europeia à Responsabilidade da Inteligência Artificial

4.1. Livro Branco sobre inteligência artificial – uma abordagem europeia virada para a excelência e confiança

A União Europeia desde muito cedo reconheceu que, com o passar dos anos, a tecnologia digital foi-se tornando um elemento cada vez mais central na sociedade e na vida das pessoas, tornando imperativo que se apoiasse uma abordagem regulamentar orientada para o seu investimento, mas nunca em detrimento da segurança e confiança dos consumidores. Assim, foi adotado pela Comissão Europeia, em fevereiro de 2020, o Livro Branco sobre Inteligência Artificial, com o objetivo de criar condições para que se conseguisse, em simultâneo, promover a adoção da Inteligência Artificial e abordar os riscos associados a determinadas utilizações destas novas tecnologias. Desta forma, o presente documento *“apresenta opções políticas que permitem um desenvolvimento fiável e seguro da inteligência artificial na Europa, no pleno respeito dos valores e dos direitos dos cidadãos europeus”* ³⁵².

O método aqui referido, é a construção de um Ecossistema de Excelência, que visa dar confiança aos cidadãos para adotarem aplicações de IA e segurança jurídica às empresas e organizações públicas para inovarem, enquanto garantem o respeito pelas regras da União, incluindo o respeito pelos direitos fundamentais e direitos dos consumidores em especial. Este ecossistema de excelência pretendia contar com contributos de inúmeras ações a vários níveis. Num primeiro momento, o plano passou pelo estabelecimento de uma cooperação mais estreita e eficaz entre Comissão e Estados-Membros de maneira a promover o desenvolvimento e a utilização da IA na Europa ³⁵³. Assim, pretendia-se a criação de vários centros de investigação europeus (no domínio da IA) de maneira que, não só se melhorasse a excelência, mas também se atraísse os melhores investigadores, o que consequentemente faria crescer as competências gerais para desenvolver e trabalhar com as tecnologias baseadas em IA ³⁵⁴. Uma vez atingido um elevado grau de especialização de matéria de IA, seria igualmente importante dar relevo às PMEs, assegurando que elas tivessem acesso e

³⁵² Livro Branco sobre inteligência artificial – Uma abordagem europeia virada para a excelência e a confiança, (COM(2020) 65 final), Comissão Europeia, Bruxelas, 2020, p.4.

³⁵³ Esse plano de cooperação propunha várias ações em domínios fundamentais como a investigação, o investimento, a entrada no mercado, as competências e o talento, os dados e a cooperação internacional. Livro Branco sobre inteligência artificial, p.6.

³⁵⁴ Também se propunha que se estipulasse e apoiasse redes de universidades e instituições de ensino superior de primeiro plano, atraindo não só os melhores professores, mas também criando uma geração com melhor preparação e competência. Livro Branco sobre inteligência artificial, p.7.

utilizassem IA ³⁵⁵. Todo este investimento seria feito em parceria com o setor privado, incentivando-se também o setor público a adotar inteligência artificial em vários domínios de interesse público ³⁵⁶.

Um outro aspeto ao qual o Livro Branco deu relevo foi ao quadro regulamentar para a IA, reconhecendo que, apesar das diversas oportunidades que a inteligência artificial apresenta, também comporta riscos para os direitos fundamentais e segurança dos cidadãos, quer seja por ter efeitos não intencionais, quer seja por poder ser utilizada para fins maliciosos. Desta forma, propõe um quadro regulamentar europeu claro para reforçar a confiança dos consumidores e das empresas em aplicações de IA, acelerando, conseqüentemente, a aceitação da tecnologia ³⁵⁷. A estratégia definida pela Comissão alinhou-se com as orientações publicadas pelo GPAN ³⁵⁸, a propósito do desenvolvimento de IA de confiança, onde foi determinado que, para se alcançar a aceitação destes sistemas de inteligência artificial, seriam necessários três componentes: a conformidade com a legislação; o respeito pelos princípios éticos e por fim a robustez. Assim, baseando-se nessas componentes e nos valores europeus, para se desenvolver sistemas de IA de confiança deverá de se verificar sete requisitos: Iniciativa e controlo por humanos; Robustez e segurança; Privacidade e governação dos dados; Transparência; Diversidade, não discriminação e equidade; Bem-estar soecietal e ambiental e por fim Responsabilização ³⁵⁹. Apesar de alguns requisitos já se encontrarem refletidos nos regimes jurídicos, outros ainda não eram abrangidos pela legislação, implicando assim mais esforços nesse sentido.

Por sua vez, são destacados alguns pontos especificamente dirigidos aos riscos que estas novas tecnologias podem comportar para os direitos fundamentais (incluindo a proteção da privacidade e dos dados pessoais e a não discriminação) bem como para a segurança e o funcionamento do regime de responsabilidade. São elencadas as características específicas destes sistemas de IA supramencionadas (como a opacidade, a complexidade, a imprevisibilidade e o comportamento autónomo) que dificultam a

³⁵⁵ Livro Branco sobre inteligência artificial, pp.8.

³⁵⁶ Quanto ao setor público, pretendia-se incentivar domínios públicos essenciais como as administrações públicas, os hospitais, os serviços de transporte e de utilidade geral pública, os supervisores financeiros. Livro Branco sobre inteligência artificial, pp.8 e 9.

³⁵⁷ Livro Branco sobre inteligência artificial, p.11.

³⁵⁸ Grupo de Peritos de Alto Nível sobre a Inteligência Artificial, “*Orientações Éticas para uma IA de Confiança*”, Comissão Europeia, Bruxelas, abril de 2018.

³⁵⁹ Para mais sobre estas orientações éticas para uma IA de confiança, consultar “*Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões - Aumentar a confiança numa inteligência artificial centrada no ser humano*”, Comissão, Bruxelas, abril de 2019; e Grupo de Peritos de Alto Nível sobre a Inteligência Artificial, “*Orientações Éticas para uma IA de Confiança*”, Comissão Europeia, Bruxelas, abril de 2018.

verificação das regras de responsabilidade em vigor na União. Neste sentido, a Comissão concluiu que o quadro legislativo da UE “*continuava, em princípio, a ser plenamente aplicável independentemente do envolvimento da IA*”, contudo, já considerava adequado melhorar alguns aspetos para fazer face a determinados riscos e situações como a aplicação e execução efetivas da legislação nacional e da União em vigor, principalmente com o cuidado que estes sistemas requerem, dadas as suas específicas características (visto que a opacidade, imprevisibilidade e complexidade da IA dificultam a identificação e prova de possíveis violações da legislação e o preenchimento das respetivas condições para se pedir uma indemnização) ³⁶⁰. Também são destacados outros riscos como a constante evolução dos sistemas de IA (que, estando sujeitos a atualizações de *software*, podem alterar o funcionamento dos produtos e até representarem novos riscos ³⁶¹), as incertezas quanto à repartição de responsabilidade entre diferentes operadores económicos na cadeia de abastecimento (por exemplo quando quem coloca o produto no mercado não é o seu produtor ³⁶²) e até mesmo às alterações que o conceito de segurança terá de sofrer (novos riscos associados a ciber ameaças ou riscos de segurança pessoal ³⁶³).

Segundo a Comissão, este novo quadro regulamentar para a IA deve de seguir uma abordagem baseada no risco “*para garantir que a intervenção regulamentar é proporcionada*” ³⁶⁴. É desta forma que somos introduzidos ao conceito de IA de “alto risco”, que deve de ser definida de forma clara e facilmente compreensível. De acordo com o Livro Branco, inteligências artificiais consideradas de alto risco serão aquelas que “*à luz do que está em causa, tendo conta a questão de saber se tanto o setor como a utilização prevista envolvem riscos significativos, nomeadamente do ponto de vista da proteção da segurança, dos direitos dos consumidores e dos direitos fundamentais*” ³⁶⁵. Portanto, para que seja considerada de alto risco terá de cumulativamente verificar que, em primeiro lugar, essa aplicação de IA é utilizada num setor em que se pode esperar que ocorram riscos significativos, pelas características das atividades tipicamente realizadas, ou seja, a utilização terá de realizar-se em domínios onde se considera mais provável a ocorrência de danos (como nos cuidados de saúde, transportes, energia e partes do setor público) ³⁶⁶; em

³⁶⁰ Livro Branco sobre inteligência artificial, p.11.

³⁶¹ *Ibidem*, p.15.

³⁶² *Ibidem*, p.16.

³⁶³ *Ibidem*.

³⁶⁴ *Ibidem*, p.19.

³⁶⁵ *Ibidem*.

³⁶⁶ Os setores em causa devem de ser enumerados de forma específica e exaustiva no novo quadro regulamentar, devendo esta lista ser revista e alterada periodicamente, sempre que necessário, de maneira que se acompanhe as respetivas evoluções. *Ibidem*.

segundo lugar, a utilização desses sistemas de IA tem de acontecer de uma maneira que é provável que surjam riscos significativos, isto pois nem sempre a utilização de IA vai implicar riscos, como por exemplo, se num dos setores anteriormente referidos, como nos cuidados de saúde, se utilizar IA como um sistema de marcação de horários de consultas, algo que, na eventualidade de uma falha do sistema, não causará nenhum risco significativo “*que justifiquem intervenção legislativa*”³⁶⁷. Uma vez definidos os tipos de IA que se consideram de alto risco, o quadro regulamentar, que estipularia requisitos obrigatórios em matéria de IA, apenas seria aplicado a estes cenários de elevado risco³⁶⁸. Excepcionalmente, podemos ter determinadas aplicações de IA que ao serem utilizadas para certos fins, devem de ser consideradas de alto risco independentemente do setor em causa, por exemplo a utilização de IA para processos de recrutamento e em situações onde afetem direitos dos trabalhadores, ou as utilizações de IA para efeitos de identificação biométrica à distância e de outras tecnologias de vigilância intrusivas, devem de ser sempre consideradas de alto risco pelo impacto que tem nos direitos das pessoas³⁶⁹.

O Livro Branco é, no fundo, um trabalho preparatório que estipulou importantes orientações e guias, principalmente com a introdução de uma abordagem baseada no risco, para promover a capacidade de inovação da Europa no domínio da IA, mas garantindo o respeito e cumprimento dos seus princípios e valores, ou seja, com os seus relatórios e conclusões visam atingir um equilíbrio entre o desenvolvimento e adoção de sistemas de inteligência artificial e uma utilização segura e fiável destas tecnologias. Foi, portanto, uma base para muitos dos documentos que se seguem, principalmente, para o Regulamento de Inteligência Artificial.

4.2. Proposta do Parlamento Europeu (2020/2014(INL))

Em outubro de 2020, deparámo-nos com uma proposta do Parlamento Europeu numa tentativa de articular um quadro legal relativo à responsabilidade civil por danos causados

³⁶⁷ Para de avaliar o nível de risco das diversas utilizações de IA, deve de se utilizar como critério o impacto nas partes afetadas, como por exemplo em casos onde estes sistemas possam produzir efeitos jurídicos ou efeitos significativos nos direitos das pessoas singulares ou empresas, ou então caso constituam um risco de lesão, morte ou danos materiais ou imateriais ou caso se produzam efeitos que não possam ser razoavelmente evitados por pessoas singulares ou coletivas. *Ibidem*.

³⁶⁸ Requisitos obrigatórios como a conservação de registo e de dados, a prestação de informações, dados usados para treinar IA em conformidade com os valores e princípios europeus, robustez e exatidão, supervisão humana e requisitos específicos para determinadas aplicações de IA como por exemplo as de identificação biométrica. Todos estes aspetos encontram-se, ao dia de hoje, previstos no Regulamento de Inteligência Artificial, tal como se fará a devida referência nos pontos que se seguem.

³⁶⁹ Livro Branco sobre inteligência artificial, p.20.

por sistemas de inteligência artificial. Esta resolução insta a Comissão a rever a Diretiva 85/374/CEE, do Conselho, de 25 de julho de 1985, para que esta apresente uma solução mais apropriada ao contexto vivido, sem detrimento de uma visão voltada para o futuro (“a Diretiva relativa à responsabilidade decorrente dos produtos (DRP) tem, há mais de 30 anos, provado ser um meio eficaz de indemnização pelos danos causados por um produto defeituoso, mas deve, no entanto, ser revista para ser adaptada ao mundo digital e dar resposta aos desafios colocados pelas tecnologias digitais emergentes, garantindo, assim, um elevado nível de proteção eficaz dos consumidores”^{370 371}).

É de destacar que esta resolução “exclui a apresentação de regras sobre o produtor, restringindo o seu âmbito ao operador de *frontend* e de *backend*”³⁷², tal como se pode observar no seu artigo 1.º, “O presente regulamento estabelece regras aplicáveis às ações de responsabilidade civil de pessoas singulares e coletivas contra operadores de sistemas de IA.”³⁷³. Deste modo, define o “operador de *frontend*” como qualquer pessoa singular ou coletiva que exerça um grau de controlo sobre um risco relacionado com a operação e funcionamento do sistema de IA e que beneficie da sua operação³⁷⁴. Por sua vez, o “operador de *backend*” é qualquer pessoa singular ou coletiva que, de forma contínua, defina as características da tecnologia, forneça dados e preste serviços essenciais de apoio de *backend* e, por conseguinte, exerça igualmente algum controlo sobre o risco ligado à operação e ao funcionamento do sistema de IA³⁷⁵. Portanto, atribui responsabilidade “aos agentes que exerçam um controlo³⁷⁶ sobre o risco de operação e de funcionamento do sistema de inteligência artificial.”³⁷⁷.

³⁷⁰ Resolução do Parlamento Europeu de 20 de outubro de 2020, que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à Inteligência Artificial, (2020/2014(INL)), considerando 9.

³⁷¹ Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, pp.249.

³⁷² Henrique Sousa Antunes, “A Responsabilidade Civil aplicável à Inteligência Artificial: Primeiras notas críticas sobre a resolução do Parlamento Europeu de 2020”, em *Revista de Direito da Responsabilidade*, ano 3, 2021, p.3.

³⁷³ (2020/2014(INL)), art. 1.º.

³⁷⁴ Cf. art. 3.º/ alínea e).

³⁷⁵ Cf. art. 3.º/ alínea f).

³⁷⁶ O controlo a que a resolução se refere é visto como qualquer ação de um operador que influencie a operação de um sistema de IA e, por conseguinte, a medida em que o operador expõe terceiros aos riscos potenciais associados à operação e ao funcionamento do sistema de IA; essas ações podem ter impacto na operação, determinando os dados introduzidos, os dados de saída ou os resultados, ou modificar funções ou processos específicos no sistema de IA; o grau em que esses aspetos da operação dos sistema de IA são determinados pela ação depende do nível de influência que o operador tem sobre o risco ligado à operação e ao funcionamento do sistema de IA. Cf. art. 3.º/ alínea g).

³⁷⁷ Henrique Sousa Antunes, “A Responsabilidade Civil aplicável à Inteligência Artificial: Primeiras notas críticas sobre a resolução do Parlamento Europeu de 2020”, p.3.

No seu artigo 2.º/n.º 1 constatamos que a proposta cobre os casos onde uma atividade, um dispositivo ou um processo físico ou virtual baseado num sistema de IA tenha causado prejuízos ou danos à vida, à saúde, à integridade física de uma pessoa singular, ao património de uma pessoa singular ou coletiva ou tenha causado danos não patrimoniais significativos que resultem numa perda económica verificável ³⁷⁸.

De seguida a resolução apresenta-nos uma listagem de definições que possam ser relevantes para a estipulação de um regime que melhor tutela a responsabilidade pelos danos causados por sistemas de inteligência artificial não só definindo “Sistema de IA” (“*um sistema baseado em software ou integrado em dispositivos físicos e que apresenta um comportamento que simula inteligência, nomeadamente recolhendo e tratando dados, analisando e interpretando o seu ambiente e tomando medidas – com um determinado nível de autonomia – para atingir objetivos específicos*” ³⁷⁹), como também contendo um conceito para “autonomia” (“*um sistema de IA que funciona interpretando certos dados e utilizando um conjunto de instruções predeterminadas, sem estar limitado a essas instruções, apesar de o comportamento do sistema estar limitado pelo objetivo que lhe foi atribuído e que está destinado a realizar e por outras escolhas de conceção tomadas por quem o desenvolveu*” ³⁸⁰).

Chegamos agora ao ponto mais característico desta resolução, a sua natureza dualista da responsabilidade, objetiva e subjetiva. Esta proposta fomentava a criação de um esquema baseado em riscos onde, a diferenciação dos vários sistemas de inteligência artificial baseado em critérios de autonomia e aleatoriedade, resultaria na aplicação de diferentes regimes de responsabilidade: para os sistemas de IA de alto risco (*High Risk AI System*) teríamos a responsabilidade objetiva, enquanto para os restantes sistemas, não considerados de alto risco, aplicaríamos uma responsabilidade subjetiva, baseada na culpa ³⁸¹. Vejamos.

³⁷⁸ Aspeto que é bastante criticado por alguns autores. Henrique Antunes mostra-se muito crítico da situação, principalmente no referente à tradução portuguesa que “*identifica estas lesões como “danos não patrimoniais”*”. *A opção é infeliz, por duas razões essenciais: sobrepondo o dano à ilicitude, parece subtrair à lesão aos demais bens considerados na norma a indemnização dos danos insuscetíveis de avaliação pecuniária; (...) a indemnização destas lesões está subordinada à verificação cumulativa de uma perda económica do lesado, a compensação de danos não patrimoniais dependeria da relevância patrimonial da lesão, esvaziando o sentido da sua automatização conceptual.*”, Henrique Sousa Antunes, “A Responsabilidade Civil aplicável à Inteligência Artificial: Primeiras notas críticas sobre a resolução do Parlamento Europeu de 2020”, p.13.

No mesmo sentido Mafalda Miranda Barbosa também menciona estes “*danos não patrimoniais que tenham um impacto económico*”, partilhando das críticas supramencionadas ao afirmar que “*parece ignorar a essência dos danos não patrimoniais, associando-os a consequências economicamente relevantes*”. Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, pp.353/354.

³⁷⁹ Cf. art. 3.º/ alínea a).

³⁸⁰ Cf. art. 3.º/ alínea b).

³⁸¹ Está aqui em causa uma tentativa de princípio de proporcionalidade, que será desenvolvido mais à frente.

O artigo 4.º/n.1 estipula que o operador de um sistema de IA tem a responsabilidade objetiva por quaisquer prejuízos ou danos causados por uma atividade, um dispositivo ou um processo físico ou virtual baseado nesse sistema de IA, desde que se trate de um sistema de inteligência artificial de alto risco. Deste modo, para se considerar um sistema de IA como um sistema de *high risk*, significa que ele tem um potencial enorme de, com a sua autonomia, causar prejuízos ou danos a uma ou a várias pessoas de uma forma aleatória e que vai para além do que se pode razoavelmente esperar³⁸². Para se medir essa potencialidade, terá de se ter em conta aspetos como a gravidade dos eventuais prejuízos ou danos, o grau de autonomia de decisão, a probabilidade de o risco se concretizar e a forma e o contexto em que o sistema de IA é utilizado³⁸³. Outro aspeto importante relacionado com esta responsabilidade objetiva é que o operador só não será considerado responsável pelos danos se estes tiverem sido causados por motivos de força maior, caso contrário, não se poderá eximir da sua responsabilidade (nem invocando que agiu com a devida diligência ou respetiva autonomia do sistema de IA³⁸⁴).

Quanto a estes sistemas de alto risco, a resolução pretendia que estes fossem enumerados de forma exaustiva num anexo ao presente regulamento e, como estamos perante uma área onde a técnica e o mercado a nível mundial evoluem a uma enorme velocidade, deveriam de se fazer revisões de 6 em 6 meses³⁸⁵ dessa listagem de sistemas de IA de forma que incluam os novos tipos de sistemas de alto risco e os setores críticos em que são implantados, suprimam os tipos de sistemas que deixem de ser considerados de alto risco ou mudem os setores críticos para os sistemas de IA de alto risco existentes³⁸⁶.

Ainda dentro das situações de danos causados por sistemas de alto risco, ainda se estipulou a contratação de um seguro de responsabilidade obrigatório para o operador, adequado aos montantes e à dimensão que estão previstos nos artigos 5.º e 6.º. Aspeto que recebeu congratulação por parte de Malfada Miranda Barbosa pelo facto de surgir “*paredes-meias com a definição de critérios de imputação baseados no risco*”, isto pois esta

³⁸² O Parlamento Europeu ao optar por este sistema baseado no risco “*trilha um caminho já aberto pela Comissão Europeia no Livro Branco*” assim é correto afirmar, que, segundo Henrique Antunes, o trabalho precedente da Comissão “*influenciou, notoriamente, a restrição da responsabilidade objetiva a sistemas de inteligência artificial de alto risco na Resolução de 2020.*”, Henrique Sousa Antunes, “*A Responsabilidade Civil aplicável à Inteligência Artificial: Primeiras notas críticas sobre a resolução do Parlamento Europeu de 2020*”, p.4.

³⁸³ Cf. art. 3.º/ alínea c).

³⁸⁴ Cf. art. 4.º/n.º 3.

³⁸⁵ Através de atos delegados nos termos do artigo 290.º do Tratado de Funcionamento da União Europeia que estabelece que um ato legislativo pode delegar na Comissão o poder de adotar atos não legislativos de alcance geral que completem ou alterem certos elementos não essenciais do ato legislativo.

³⁸⁶ Cf. art. 4.º/n.º 2, alíneas a), b) e c).

responsabilidade objetiva “*não se queda numa pura responsabilidade pela causalidade (...), antes se configura como uma responsabilidade assente no risco*”³⁸⁷. Estes seguros podem ajudar a garantir que as vítimas recebem uma indemnização efetiva e a agregar os riscos de todos os segurados^{388 389}.

Presente no artigo 5.º temos as limitações aos danos indemnizáveis sendo o montante máximo de dois milhões de EUR em caso de morte ou de danos causados à saúde ou à integridade física de uma pessoa lesada em resultado de uma operação de um sistema de IA de alto risco e o montante máximo de um milhão de EUR em caso de danos não patrimoniais significativos que resultem numa perda económica verificável ou de danos causados ao património, incluindo quando vários bens de uma pessoa lesada ficarem danificados em resultado de uma única operação de um mesmo sistema de IA de alto risco. Mais, a resolução prevê a possibilidade da criação de fundos de compensação³⁹⁰ que sirvam para colmatar lacunas em matéria de seguro, como por exemplo nas situações onde ocorram danos coletivos, em que a indemnização ultrapasse significativamente os montantes máximos estabelecidos (situações em que os limites estabelecidos afiguram-se “*manifestamente insuficientes*”³⁹¹), ou também nos casos excecionais em que um sistema de IA ainda não classificado como um sistema de IA de alto risco e, por conseguinte, ainda não segurado, provoque prejuízos ou danos³⁹².

Por outro lado, sempre que estivermos perante sistemas de IA que não são considerados de alto risco nos termos mencionados anteriormente, teremos a aplicação da responsabilidade subjetiva (“O operador de um sistema de IA que não constitua um sistema de IA de alto risco (...) está sujeito à responsabilidade culposa por quaisquer prejuízos ou danos causados”³⁹³). Podemos dizer que esta é a verdadeira regra desta resolução de 2020

³⁸⁷ Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.352.

³⁸⁸ Considerando 21.

³⁸⁹ Um aspeto importante a considerar é que, no início, reconhece-se que haja alguma dificuldade na constituição destes seguros pela falta de acesso a dados de alta qualidade, ou devido a quantidades insuficientes de dados referentes a estas novas tecnologias, já que um dos fatores nos quais as seguradoras baseiam a sua oferta de produtos e serviços é a avaliação dos riscos baseada no acesso a dados históricos suficientes relativos a ações. Por isso, para se conseguir uma cobertura mais inovadora terá de haver mais acesso aos dados, por exemplo através da estipulação de uma obrigação de fornecer informações bem documentadas. (Considerando 21).

³⁹⁰ Fundos de carácter subsidiário, o que no entender de Mafalda Miranda Barbosa é “*preferível*”. Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.352.

³⁹¹ *Ibidem*, p.353.

³⁹² Considerando 22.

³⁹³ Cf. art. 8.º/n.º 1.

se tivermos em conta que o nível atual da autonomia ainda não conheceu a evolução que se prognostica para o futuro ³⁹⁴.

Diversamente da regra nos regimes tradicionais, aqui temos uma presunção de culpa ilidível por parte do operador (inversão do ónus da prova), pois será este que pode afastar a responsabilidade se provar que os danos ou prejuízos foram causados sem culpa da sua parte, baseando-se num dos seguintes motivos: primeiro, provar que o sistema de IA foi ativado sem o seu conhecimento, embora tenham sido tomadas todas as medidas razoáveis e necessárias para evitar essa ativação fora do controlo do operador; ou em segundo lugar, provar que foi observada a devida diligência através da execução de várias ações como a seleção de um sistema de IA adequado para as tarefas e capacidades em causa, a correta colocação em operação do sistema de IA, o controlo das atividades e manutenção da fiabilidade da operação, graças à instalação regular de todas as atualizações disponíveis ³⁹⁵. Não conseguindo afastar esta presunção de culpa, o operador responderá pelos danos causados. Apesar disto, mais uma vez, o operador não é considerado responsável se os danos forem causados por motivos de força maior, não podendo este furtar-se à responsabilidade invocando que os danos foram causados por uma atividade, um dispositivo ou um processo autónomo baseado no seu sistema de IA.

Para além desta responsabilidade subjetiva implicar uma presunção de culpa, surge agravada pelo facto de impor que o operador seja responsável pelo pagamento da indemnização caso os danos ou prejuízos tenham sido causados por um terceiro que interferiu no sistema de IA, alterando o seu funcionamento ou efeitos, mas apenas quando este terceiro não for localizável ou carecer de recursos financeiros ³⁹⁶. Trata-se de uma “*responsabilidade objetiva, por facto alheio*” podendo ainda existir uma responsabilização geral do operador pela violação de deveres de cuidado, caso este não tenha evitado a referida interferência.

Olhando de uma maneira geral, esta proposta realizada pelo Parlamento sofre de um problema, falta de ambição. No fundo, esta resolução “*is nothing more than an adaptation, in the sense of transforming and updating the Defective Products Directive (...). It cannot be said to have been ambitious or innovative in the solutions proposed in the resolution for*

³⁹⁴ Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.355.

³⁹⁵ Cf. art. 8.º/n.º 2. Também Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, p.250.

³⁹⁶ Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.351/352.

liability for damage caused by AI systems, solutions that will certainly fall short of the ecosystem of excellence and trust sought in the field of AI.”³⁹⁷. Também Henrique Sousa Antunes apura a falta de ambição como uma crítica ao proposto pelo Parlamento Europeu, embora esta sua visão seja mais direcionada ao facto de a resolução de 2020 ter abandonado a via da atribuição de uma personalidade eletrónica aos entes dotados de IA (algo que já tinha sido levantado na resolução de 2017)³⁹⁸. Aqui, mantemos a posição partilhada anteriormente, de rejeitar o caminho da atribuição de uma personalidade eletrónica, pelo menos neste “temporário” estado inicial, pela não verificação de condições que exijam uma igualdade de dignidade entre seres humanos e máquinas.

4.3. Regulamento de Inteligência Artificial

A Comissão Europeia, a 21 de abril de 2021, apresentou ao Parlamento Europeu e ao Conselho, uma proposta de Regulamento que visava criar regras harmonizadas em matéria de inteligência artificial³⁹⁹. Esta proposta deu origem ao Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024, doravante chamado de Regulamento de Inteligência Artificial (ou por referência à sigla “RIA”)⁴⁰⁰.

O objetivo deste regulamento é melhorar o funcionamento do mercado interno mediante a previsão de um regime jurídico uniforme em matéria de IA que, em particular, promova o desenvolvimento e regule a colocação no mercado, a colocação em serviço e a utilização de sistemas de inteligência artificial na União, em conformidade com os valores desta, a fim de promover a adoção de uma inteligência artificial centrada no ser humano e de confiança. Desta maneira pretende-se assegurar, simultaneamente, um elevado nível de proteção de interesses públicos, como a saúde, a segurança e os direitos fundamentais consagrados na Carta dos Direitos Fundamentais da UE (nomeadamente a democracia, o

³⁹⁷ Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, p.251.

³⁹⁸ Henrique Sousa Antunes, “A Responsabilidade Civil aplicável à Inteligência Artificial: Primeiras notas críticas sobre a resolução do Parlamento Europeu de 2020”, p.17 e 18.

³⁹⁹ Ao abrigo do artigo 294.º do Tratado de Funcionamento da União Europeia.

⁴⁰⁰ O Regulamento entrou em vigor no passado dia 1 de agosto de 2024 (“entre em vigor no vigésimo dia seguinte ao da sua publicação no Jornal Oficial da União Europeia”, que ocorreu a 12 de julho de 2024) Cf. art. 113.º do RIA.

Estado de direito e a proteção do ambiente), bem como assegurar uma proteção contra efeitos nocivos dos sistemas de IA na União enquanto se apoia a inovação ⁴⁰¹.

Este era um esforço necessário para se assegurar um nível elevado de proteção compatível com o desenvolvimento de IAs de confiança, ao mesmo tempo que se evitavam divergências que prejudicassem a livre circulação, inovação, implantação e adoção dos sistemas de IA e dos restantes produtos e serviços conexos no mercado interno. Assim, estipulam-se regras uniformes para que se acautele diferenças entre as regras nacionais, que podem conduzir à fragmentação do mercado interno e reduzir a segurança jurídica para os operadores que desenvolvem, importam ou utilizam sistemas de IA ⁴⁰².

Já que o regulamento pretende uma harmonização das regras relativas a sistemas de IA, de maneira a promover um desenvolvimento e utilização segura e compatível com os valores da União Europeia, não podia deixar de fazer menção a uma definição de Sistema de inteligência artificial. Ora, de acordo com o artigo 3.º - 1), “Sistema de IA, é um sistema baseado em máquinas concebido para funcionar com níveis de autonomia variáveis, e que pode apresentar capacidade de adaptação após a implantação e que, para objetivos explícitos ou implícitos, e com base nos dados de entrada que recebe, infere a forma de gerar resultados, tais como previsões, conteúdos, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais”. Deste conceito, podemos concluir que o Regulamento de Inteligência Artificial, apresenta-nos uma noção um pouco ampla, abordagem que parece ser intencional de maneira que consiga preconizar situações futuras ⁴⁰³. É ainda notória a importante menção expressa ao conceito de autonomia, tal como referida anteriormente a propósito da Resolução do Parlamento Europeu de 2020, com a particularidade de referir a possibilidade de vários níveis de autonomia.

É pertinente referir que o Regulamento expressamente exclui do seu âmbito sistemas de IA que tenham sido colocados no mercado, colocados em serviço ou utilizados exclusivamente para finalidades militares, de defesa ou de segurança nacional, bem como os sistemas que, embora não sejam colocados no mercado ou colocados em serviço na União, os seus resultados forem utilizados pela União com as mesmas finalidades ⁴⁰⁴. Do mesmo

⁴⁰¹ Considerando 1 do RIA; Cf. art. 1.º.

⁴⁰² Considerando 3 do RIA.

⁴⁰³ “It provides for a technology-neutral definition of AI systems, which is already future-oriented and may encompass techniques and approaches that are not yet developed or known”, Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, p.251.

⁴⁰⁴ Cf. art. 2.º/n.º 3.

modo se exclui a aplicação do Regulamento a sistemas de IA ou modelos de IA, incluindo os respetivos resultados, especificamente desenvolvidos e colocados em serviço exclusivamente para fins de investigação e desenvolvimento científicos ⁴⁰⁵. São ainda previstas mais algumas situações excluídas do âmbito do regulamento.

A abordagem que é verificada no Regulamento de Inteligência Artificial é baseada num sistema de risco repartido em diferentes níveis: sistemas de IA de riscos inaceitáveis, sendo por isto proibidos; sistemas de IA de elevado risco, aos quais o Regulamento estabelece requisitos específicos e rigorosas obrigações para os seus operadores; sistemas de IA de risco limitado e por fim aqueles que apresentem um risco reduzido ou nenhum risco, relativamente aos quais não se verificam especiais exigências ⁴⁰⁶. Há ainda uma menção especial aos modelos de inteligência artificial de finalidade geral aos quais o Regulamento presta algumas considerações. Vejamos as especificidades de cada nível.

Em primeiro lugar, temos a menção aos sistemas de inteligência artificial que comportem um risco inaceitável, presentes no artigo 5.º sob a epígrafe “Práticas de IA proibidas”. Estamos a falar das situações onde sistemas de inteligência artificial podem ser utilizados indevidamente e conceder instrumentos novos e poderosos para práticas manipuladoras, exploratórias e de controlo social, ou seja, são atividades consideradas prejudiciais e abusivas por desrespeitarem os valores da União (como a dignidade do ser humano, a liberdade, a igualdade, a democracia e o Estado de direito bem como os direitos fundamentais como o direito à não discriminação, à proteção de dados pessoais e à privacidade e os direitos das crianças), devendo por isto mesmo, ser proibidas ⁴⁰⁷. Assim, estão proibidas práticas como a colocação no mercado, a colocação em serviço ou a utilização de sistemas de IA que tentem manipular ou controlar os comportamentos de uma

⁴⁰⁵ Cf. art. 2.º/n.º 6.

⁴⁰⁶ Susana Aires de Sousa, “Breves Notas sobre a “Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União”, em *Direito em Mudança: A Proposta de Regulamento Europeu sobre Inteligência Artificial – Algumas Questões Jurídicas*, Instituto Jurídico, Faculdade de Coimbra, Universidade de Coimbra, julho, 2023, p.8; Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.358.

⁴⁰⁷ Considerando 28.

ou mais pessoas ⁴⁰⁸, que explorem as vulnerabilidades das pessoas singulares ⁴⁰⁹, procedam a avaliações ou classificações de vários indivíduos e respetivos grupos ⁴¹⁰ ou realizem avaliações de risco ⁴¹¹. Ainda dentro dos sistemas proibidos também encontramos aqueles que são utilizados para criar ou expandir bases de dados de reconhecimento facial através da recolha aleatória de imagens faciais a partir da internet ou de imagens de televisão em circuito fechado (TVCF) ⁴¹², bem como os sistemas utilizados para inferir emoções de uma de uma pessoa singular no local de trabalho e nas instalações de ensino ⁴¹³.

São também estipuladas várias proibições a sistemas voltados para uma vertente biométrica, como aqueles que realizem categorizações biométricas ⁴¹⁴, ou que utilizem identificação biométrica à distância em “tempo real” em espaços acessíveis ao público para efeitos de aplicação da lei (apesar desta proibição, temos exceções quando se prosseguem fins específicos ⁴¹⁵).

⁴⁰⁸ Sistemas de IA que empreguem técnicas subliminares ou não, que contornem a consciência de uma pessoa, ou que tentam manipulá-la com o objetivo de distorcer substancialmente o seu comportamento, prejudicando consideravelmente a sua capacidade de tomar decisões, ou fazendo com que tome decisões que, caso contrário, não tomaria, havendo a probabilidade de provocar, ou chegando mesmo a provocar danos significativos (repercussões negativas suficientemente importantes na saúde física, psicológica ou nos interesses financeiros). Portanto, prejudicam a sua autonomia, tomada de decisões e liberdade de escolha através de estímulos que ultrapassam a perceção humana (como estímulos de áudio, de imagem e de vídeo), que para além de subverterem o comportamento humano, fazem-no de maneira que as pessoas não tenham consciência dessas mesmas técnicas. Isto pode acontecer, por exemplo, através de máquinas-cérebro ou por realidade virtual, onde há grande controlo dos estímulos a que as pessoas são expostas. Cf. art. 5.º/n.º 1-a); Considerando 29.

⁴⁰⁹ Proibidos os sistemas que explorem as vulnerabilidades de uma pessoa ou de um grupo específico de pessoas devidas à sua idade, incapacidade ou situação socioeconómica, provocando danos significativos. A idade, incapacidade ou situação socioeconómica tornam as pessoas mais vulneráveis à exploração, como aqueles que vivem em pobreza extrema ou minorias étnicas ou religiosas. Cf. art. 5.º/n.º 1-b); Considerando 29.

⁴¹⁰ Sistemas de IA que avaliem ou classifiquem pessoas singulares ou grupos de pessoas, baseando-se nos seus comportamentos sociais ou características pessoais (conhecidas, inferidas ou previsíveis). Estas classificações sociais podem conduzir a tratamentos prejudiciais ou desfavoráveis das pessoas alvo em contextos sociais não relacionados com os contextos nos quais os dados foram originalmente gerados ou recolhidos, ou levarem a tratamentos injustificados ou desproporcionados face ao seu comportamento social ou à gravidade do mesmo. Cf. art. 5.º/n.º 1-c) – i) e ii).

⁴¹¹ Sistemas que, com o objetivo de avaliar ou prever o risco de uma pessoa singular cometer uma infração penal, com base exclusivamente na definição de perfis ou na avaliação dos seus traços e características de personalidade. Contudo, é permitido a utilização de sistemas de IA para apoiar a avaliação humana do envolvimento de uma pessoa numa atividade criminosa, que já se baseia em factos objetivos e verificáveis diretamente ligados a uma atividade criminosa. Cf. art. 5.º/n.º 1-d).

⁴¹² Cf. art. 5.º/n.º 1-e).

⁴¹³ Excetuando-se os casos em que esse sistema de IA se destine a ser introduzido ou instalado no mercado por razões médicas ou de segurança. Cf. art. 5.º/n.º 1-f).

⁴¹⁴ Sistemas que categorizam ou classifiquem biométrica e individualmente as pessoas com base nos seus dados biométricos para deduzir ou inferir a sua raça, opiniões políticas, filiação sindical, convicções religiosas, vida sexual ou orientação sexual, excetuando-se desta proibição as rotulagens ou filtragens de conjuntos de dados biométricos legalmente adquiridos, tais como imagens, com base em dados biométricos ou na categorização de dados biométricos no domínio da aplicação da lei. Cf. art. 5.º/n.º 1-g).

⁴¹⁵ A utilização destes sistemas de identificação biométrica à distância em “tempo real” é possível na medida em que essa utilização é estritamente necessária para fins de: busca seletiva de vítimas específicas de rapto, tráfico de seres humanos ou exploração sexual de seres humanos, bem como a busca por pessoas desaparecidas; prevenção de uma ameaça específica substancial e iminente à vida ou à segurança física de pessoas singulares ou de uma ameaça real e atual ou real e previsível de um ataque terrorista; ou para a localização ou identificação

Em segundo lugar, temos os sistemas de IA de elevado risco com as respetivas regras de classificação previstas no artigo 6.º. Assim estaremos perante sistemas de IA de elevado risco quando se preencham duas condições: o sistema destina-se a ser utilizado como uma componente de segurança de um produto, ou então o próprio sistema de IA é o produto (um dos produtos que esteja abrangido pelo atos enumerados na lista de legislação harmonizada constante do Anexo I); depois este produto tem de ser objeto de um procedimento de avaliação de conformidade por terceiros com vista à sua colocação no mercado ou colocação em serviço (nos termos enumerados na lista de legislação harmonizada constante do Anexo I)^{416 417}.

Para além do parâmetro de classificação presente no n.º 1 do artigo 6.º, o Regulamento também estipula uma lista, presente no Anexo III, onde constam vários domínios onde a utilização de sistemas de IA deve de ser considerada de elevado risco por serem casos particularmente sensíveis. Para que a utilização de um sistema de IA num dos domínios que constam do Anexo III, resulte na classificação do sistema como sistema de elevado risco, este terá de representar um risco significativo de danos para a saúde, a segurança ou para os direitos fundamentais das pessoas singulares, por exemplo se influenciar de forma significativa o resultado das suas tomadas de decisões, caso contrário, não será considerado um sistema de elevado risco⁴¹⁸ (só nos casos em que os sistemas de IA executem definições de perfis de pessoas singulares é que serão sempre considerados de risco elevado, mesmo que não representem aquele risco significativo⁴¹⁹). Por outras palavras, este segundo parâmetro, também se apresenta com duas condições: o sistema de IA tem de

de uma pessoa suspeita de ter cometido uma infração penal, para efeitos da realização de uma investigação criminal, ou instauração de ação penal ou execução de uma ação penal por algumas das infrações referidas no Anexo II (como terrorismo, tráfico de seres humanos, exploração sexual de crianças e pornografia infantil, homicídio, ofensas corporais graves, rapto, sequestro, tomada de reféns, violação, sabotagem, criminalidade ambiental, entre outras.) Cf. art. 5.º/n.º 1-h), alíneas i), ii) e iii), juntamente com o Anexo.

⁴¹⁶ Cf. art. 6.º/n.º 1-a) e b).

⁴¹⁷ Estes produtos podem ser máquinas, brinquedos, ascensores, aparelhos e sistemas de proteção destinados a serem utilizados em atmosferas potencialmente explosivas, equipamentos de rádio, equipamentos sob pressão, equipamentos de embarcações de recreio, instalação por cabo, aparelhos a gás, dispositivos médicos, dispositivos médicos para diagnóstico *in vitro*, automóveis e aviação. (Considerando 50).

⁴¹⁸ A contrário sensu do 1.º paragrafo do n.º 3 do artigo 6.º.

De seguida, o artigo procede com uma enumeração de condições onde alguns sistemas não serão considerados de elevado risco, como quando estes desempenhem tarefas processuais restritas (por exemplo transformar dados não estruturados em dados estruturados), ou se destinem a melhorar resultados de atividades humanas previamente concluídas, se destinem a detetar padrões de tomada de decisões ou desvios em relação a padrões de tomada de decisões anteriores e não se destinem a substituir nem influenciar uma avaliação humana previamente concluída, sem que se proceda a uma verificação adequada por um ser humano, ou por fim se executarem tarefas preparatórias no contexto de uma avaliação pertinente para efeitos dos casos de utilização enumerados no Anexo III. O Considerando 53 do RIA explora mais estas condições, facultando também alguns exemplos práticos.

⁴¹⁹ Último paragrafo do n.º 3 do artigo 6.º.

representar um risco elevado de danos para a saúde e a segurança, ou de prejuízo para os direitos fundamentais das pessoas e, cumulativamente, terá de ser utilizado dentro das concretas situações de um dos oito domínios predefinidos no Anexo III ⁴²⁰.

Os domínios que se encontram enumerados no Anexo III são os seguintes: categoria especial referente aos dados biométricos ⁴²¹; gestão e funcionamento de infraestruturas críticas ⁴²²; educação e formação profissional ⁴²³; emprego, gestão de trabalhadores e acesso ao emprego por contra própria ⁴²⁴; acesso a serviços privados e públicos essenciais, bem

⁴²⁰ Considerando 52.

⁴²¹ Quanto à relação biometria e sistemas de inteligência artificial, aqui teremos como sistemas de elevado risco os sistemas de identificação biométrica à distância (não sendo de risco elevado se o único propósito é confirmar que uma determinada pessoa singular é a pessoa que alega ser), os sistemas de categorização biométrica, de acordo com atributos ou características sensíveis (na medida em que não sejam proibidos nos termos do artigo 5.º), e por fim os sistemas de reconhecimento de emoções. Considerados sistemas de elevado risco pelo perigo de resultados enviesados, com base na idade, etnia, raça, sexo ou deficiência, que podem levar a tratamentos discriminatórios. (Considerando 54; ponto 1 do anexo III).

⁴²² Aqui são classificados como sistemas de risco elevado, os sistemas de IA que se destinam a ser utilizados como componentes de segurança na gestão e no funcionamento das infraestruturas digitais críticas (enumeradas no ponto 8 do anexo I da Diretiva (EU) 2022/2557, como por exemplo ponto de troca de tráfego, serviços de DNS, redes públicas de comunicações eletrónicas, redes de distribuição de conteúdo, entre outros), do trânsito rodoviário e das redes de abastecimento de água, gás, aquecimento e eletricidade. São considerados sistemas de risco elevado, pois anomalias ou falhas destes sistemas de IA, podem resultar em perturbações substanciais das atividades sociais e económicas, bem como pôr em risco a vida e a saúde das pessoas em larga escala. Por exemplo, sistemas que monitorizem a pressão da água ou controlem os alarmes de incêndio em centros de computação em nuvem. (Considerando 55; ponto 2 do anexo III).

⁴²³ No domínio da educação podem-se colher vastas vantagens com o recurso a sistemas de inteligência artificial, contudo, estes sistemas podem ter um papel preponderante na vida de uma pessoa, determinando o seu percurso educativo e profissional o que, conseqüentemente, afeta a sua capacidade de assegurar a sua subsistência. Mais, perante falhas, podem ser particularmente intrusivos e violar o direito à educação e formação, principalmente se estiverem em causa sistemas que possam determinar o acesso ou admissão das pessoas a instituições ou programas de ensino, sistemas que avaliem os resultados de aprendizagem das pessoas, sistemas que averiguem o grau de ensino adequado para uma pessoa e influenciem os níveis de ensino e formação que recebam, bem como sistemas utilizados para monitorizar e detetar comportamentos proibidos de estudantes durante os testes. (Considerando 56; ponto 3 do anexo III).

⁴²⁴ Nomeadamente sistemas que são utilizados no recrutamento ou seleção de pessoas em vagas (através da colocação de anúncios direcionados, filtragem de candidaturas ou avaliação de candidatos), e sistemas que tomem decisões que possam afetar os termos das relações de trabalho, a promoção ou cessação das relações contratuais de trabalho, na atribuição de tarefas com base em comportamentos individuais ou outros traços característicos, e também que avaliem desempenhos e condutas das pessoas. Mais uma vez, o elevado risco justifica-se pelo impacto significativo nas perspetivas de carreira, nas subsistências das pessoas e respetivos direitos dos trabalhadores que estes sistemas podem ter. (Considerando 57; ponto 4 do anexo II).

como o usufruto dos mesmos ⁴²⁵; aplicação da lei ⁴²⁶; gestão da migração, do asilo e do controlo das fronteiras ⁴²⁷ e por fim a administração da justiça e processos democráticos ⁴²⁸.

⁴²⁵ Um domínio que também merece especial atenção é o acesso a certos serviços e prestações essenciais, tanto de cariz privado como público, e o usufruto dos mesmos. A utilização de sistemas de IA neste domínio pode ter significativos impactos nas participações de variadas pessoas na sociedade, bem como afetar o seu nível de vida. Por exemplo, sistemas de IA utilizados por autoridades públicas para avaliar a elegibilidade de pessoas singulares para terem acesso a prestações e serviços de assistência pública essenciais (como cuidados de saúde, prestações de segurança social, serviços sociais que prestam proteção em casos como maternidade, doença, acidentes de trabalho, dependência ou velhice, entre outros) determinando assim a sua concessão, recusa, redução, revogação ou recuperação, têm de ser considerados sistemas de elevado risco pelo grande impacto que podem ter na subsistência de uma pessoa, podendo até violar os seus direitos fundamentais como o direito à proteção social, à não discriminação, à dignidade do ser humano ou à ação.

Outros sistemas como aqueles que são utilizados para avaliar a capacidade de solvabilidade de pessoas singulares ou estabelecer a sua classificação de crédito (com exceção dos utilizados para deteção de fraude financeira) também comportam um risco elevado por determinarem o acesso das pessoas a recursos financeiros ou a serviços essenciais como o alojamento, a eletricidade e os serviços de telecomunicações.

Os sistemas que façam avaliações dos riscos e fixação de preços para seguros de saúde e de vida, também podem ter um forte impacto na vida das pessoas e, caso sejam mal concebidos, violar direitos fundamentais e ter consequências graves para a vida e para a saúde das pessoas.

Por fim, temos também como sistemas de elevado risco os sistemas que avaliem e classifiquem chamadas de emergência, ou enviem ou estabeleçam prioridades no envio dos serviços de primeira resposta a emergências (como polícia, bombeiros e assistência médica), bem como sistemas de triagem de doentes para cuidados de saúde de emergência, pelo facto de tomarem decisões que podem afetar a vida, saúde e os bens de uma pessoa. (Considerando 58; ponto 5 do anexo III).

⁴²⁶ No domínio da justiça, uma utilização de sistemas de IA, dentro dos termos do direito da União e do direito nacional, pode ter várias finalidades como avaliar o risco de uma pessoa singular vir a ser vítima de infrações penais, avaliar os traços e características da personalidade ou comportamento criminal, avaliar o risco de alguém cometer um infração penal ou o risco de reincidência, utilizar inteligência artificial como polígrafos ou instrumentos semelhantes, avaliar a fiabilidade dos elementos de prova ou definir o perfil de uma pessoa singular no decurso da deteção, investigação ou ação penal relativas a infrações penais. Todas estas utilizações de sistemas de IA devem de ser consideradas de risco elevado pela repercussão negativa que se pode sentir nos direitos fundamentais caso estes sistemas não sejam treinados com dados de alta qualidade, não cumpram os requisitos adequados em termos de desempenho, exatidão ou solidez e caso não tenham sido bem concebidos ou testados antes da colocação no mercado, o que podia implicar a violação de direitos fundamentais processuais como o direito à ação e a um tribunal imparcial, o direito à defesa e a presunção de inocência. (Considerando 59; ponto 6 do anexo III).

⁴²⁷ Os sistemas serão considerados de elevado risco, no domínio da gestão da migração, do asilo e do controlo das fronteiras, pois poderão afetar pessoas que se encontrem em posições de alguma vulnerabilidade, ou seja, é importante garantir que não haja por parte de sistemas de IA decisões tomadas na base de práticas discriminatórias para que não se lesem direitos à livre circulação, à não discriminação, à proteção da vida privada e dos dados pessoais, à proteção internacional e a uma boa administração. Estamos assim a falar de sistemas de IA que possam ser utilizados como polígrafos ou instrumentos semelhantes, sistemas utilizados para avaliar os riscos colocados por pessoas que entram no território de um Estado-Membro (incluindo um risco para a segurança, um risco de migração irregular ou um risco para a saúde), sistemas que sirvam de auxílio às autoridades competentes nas análises de pedidos de asilo, de visto e de autorização de residência ou sistemas que são utilizados para efeito de deteção, reconhecimento ou identificação de pessoas singulares (com a exceção da verificação de documento de viagem). (Considerando 60; ponto 7 do anexo III).

⁴²⁸ Considerados de elevado risco, pelo impacto potencialmente significativo que podem ter na democracia, no Estado de direito e nas liberdades individuais, bem como no direito à ação penal e a um tribunal imparcial. Assim, considerando potenciais anomalias, riscos de potenciais enviesamentos e opacidades são sistemas de elevado risco aqueles utilizados na investigação e interpretação de facto e do direito e na aplicação da lei, por autoridades judiciárias, ou por quem o faça em seu nome. Também encontramos aqui sistemas utilizados para resolução alternativa de litígios pelas entidades competentes, reforçando o Regulamento que, estas ferramentas de IA devem de apenas auxiliar o poder de tomada de decisão dos magistrados sem substituir a atividade humana. (Considerando 61).

Aqui também estão os sistemas concebidos para serem utilizados para influenciar o resultado de uma eleição ou referendo, ou o comportamento eleitoral de pessoas singulares no exercício do seu direito de voto

Importante mencionar que esta lista de sistemas de elevado risco do Anexo III pode sofrer alterações de acordo com o artigo 7.º do Regulamento de Inteligência Artificial. No sentido de adicionar ou modificar casos de utilização de sistemas de IA ao anexo III, a Comissão pode adotar atos delegados nos termos do artigo 97.º, desde que verifique, cumulativamente, que os sistemas de IA destinam-se a ser utilizados em qualquer um dos domínios supramencionados e, desde que esses sistemas representem um risco de danos ou repercussões negativas nos direitos fundamentais e esse risco seja superior aos riscos de danos representado pelos sistemas de IA de elevado risco já referidos pelo anexo. Para avaliar se os sistemas de IA representam ou não danos para a saúde e segurança ou repercussões negativas para os direitos fundamentais, a Comissão terá de ter em conta vários critérios enumerados no n.º 2 do artigo 7.º⁴²⁹. Por outro lado, também podemos ter a retirada de sistemas que já constam do anexo III, desde que se verifique, cumulativamente, que esse sistema deixa de representar um risco significativo para os direitos fundamentais, saúde ou a segurança (tendo em consideração os mesmos critérios do n.º 2) e desde que essa supressão não diminua o nível geral de proteção dos direitos fundamentais ao abrigo do direito da União⁴³⁰.

Um sistema ser considerado de elevado risco implica a sujeição a um conjunto vasto de obrigações e requisitos exigentes que devem de ser cumpridos sob pena de sanções⁴³¹.

O regulamento estipula que em relação aos sistemas de IA de risco elevado, deve de ser criado um sistema de gestão de riscos, que terá como objetivo identificar e atenuar os riscos de danos significativos que estes sistemas representam para a saúde, a segurança e os direitos fundamentais⁴³². Este sistema é entendido como um processo iterativo contínuo, planeado e executado ao longo de todo o ciclo de vida de um sistema de IA de risco elevado, sendo sujeito a revisões e atualizações sistemáticas regulares. É um processo composto por várias etapas onde se irá identificar e analisar os riscos conhecidos e razoavelmente previsíveis que o sistema de IA possa representar, bem como realizar-se estimativas e avaliações de riscos que possam surgir da utilização em conformidade com a sua finalidade prevista ou que possam surgir com base na análise dos dados recolhidos por meio do sistema

(excetuando os que sejam utilizados como instrumentos para organizar, otimizar e estruturar campanhas políticas de um ponto de vista administrativo e logístico). (Considerando 62; ponto 8 do anexo III).

⁴²⁹ Ter em conta as finalidades previstas no sistema de IA, o seu grau/probabilidade de utilização, a sua natureza e a quantidade dos dados tratados e utilizados pelo sistema, o seu grau de autonomia, entre outros. Cf. art. 7.º/n.º 2.

⁴³⁰ Cf. art. 7.º/n.º 3-a) e b).

⁴³¹ Que se encontram regulamentadas no artigo 99.º e seguintes do RIA.

⁴³² Considerando 65.

de acompanhamento pós-comercialização (artigo 72.º), para que se consiga assim adotar medidas adequadas e específicas de gestão de riscos concebidas para atenuar ou eliminar os que sejam previamente identificados ⁴³³.

São também estabelecidos mais alguns requisitos e regras iniciais para os sistemas de risco elevado como alguma documentação técnica que deve de ser elaborada (artigo 11.º) e manutenção de registos (artigo 12.º). Também devem de ser respeitadas regras de transparência e prestação de informações aos responsáveis pela implantação, de maneira a assegurar que o seu funcionamento seja suficientemente transparente para que aqueles responsáveis possam interpretar os resultados dos sistemas e utilizá-los corretamente. Estes sistemas devem de ser acompanhados de instruções de utilização, em formato digital ou outro adequado, que contenha informações concisas, completas, corretas e claras que sejam pertinentes, acessíveis e compreensíveis (artigo 13.º) ⁴³⁴.

O artigo 14.º impõe que os sistemas de elevado risco sejam concebidos e desenvolvidos de maneira que possam ser eficazmente supervisionados por pessoas singulares durante o período em que estão a ser utilizados (por meio de ferramentas de interface homem-máquina apropriados), com vista a prevenir ou minimizar os riscos que estes sistemas representem para a saúde, segurança e os direitos fundamentais.

Com o artigo 16.º são introduzidas várias obrigações dos prestadores de sistemas de inteligência artificial de risco elevado, como assegurar que os seus sistemas cumpram os requisitos anteriormente mencionados; indicar os seus dados de identificação e contacto no respetivo sistema, embalagem ou documentação que o acompanha; dispor de um sistema de gestão de qualidade (de acordo com o artigo 17.º) ⁴³⁵; conservar a documentação apropriada (artigo 18.º) bem como manter os registos gerados automaticamente pelos sistemas de IA (artigo 19.º), entre muitas outras obrigações. Há também obrigações dos importadores (artigo 23.º) e dos distribuidores (artigo 24.º), bem como regras para os mandatários dos prestadores (artigo 22.º). Temos ainda regulamentação face a normas harmonizadas e produtos de normalização (artigo 40.º), avaliação da conformidade (artigo 42.º e 43.º), certificados (artigo 44.º) e registo (artigo 49.º).

⁴³³ Cf. art. 9.º/n.º 2; Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.360.

⁴³⁴ Estas instruções de utilização devem incluir informações como a identidade e dados de contacto do prestador do sistema (e mandatário se for caso disso), as características, capacidades e limitações de desempenho, recursos computacionais e de *hardware* necessários, vida útil esperada do sistema, medidas de manutenção e assistência necessárias, entre outros. Cf. art. 13.º/n.º 3.

⁴³⁵ Os prestadores devem de criar um sistema de gestão de qualidade, documentado de maneira sistemática e ordenada sob forma de políticas, procedimentos e instruções escritas que assegurem a conformidade com o presente regulamento.

No terceiro nível da hierarquia de riscos temos os sistemas de IA de risco limitado que implicam obrigações de transparência aplicáveis aos prestadores e responsáveis pela implantação dos sistemas, promovendo-se assim uma maior confiança da sociedade. De acordo com o artigo 50.º, estando em causa sistemas de inteligência artificial destinados a interagir com pessoas singulares, os prestadores devem assegurar que as pessoas sabem que estão a lidar com um sistema de IA, ou seja, devem de ser informadas de que estão a interagir com máquinas, salvo se tal for óbvio do ponto de vista de uma pessoa razoavelmente informada, atenta e advertida, considerando as circunstâncias e o contexto de utilização ⁴³⁶. Do mesmo modo, também nos conteúdos sintéticos de áudio, imagem, vídeo ou texto (gerados pelos sistemas de IA), deve constar a informação de que foram artificialmente criados ou manipulados, isto porque cada vez se torna mais difícil de distinguir esses conteúdos artificiais dos gerados por seres humanos ⁴³⁷.

Esta obrigação de transparência também se aplica aos responsáveis pela implantação de um sistema de reconhecimento de emoções ou de um sistema de categorização biométrica, onde aqueles devem de informar as pessoas expostas a esse sistema do seu funcionamento e tratar os dados pessoais em conformidade com a legislação uniformizada pelo regulamento.

O regulamento ainda nos fala de modelos de IA de finalidade geral, previstos nos artigos 51.º e seguintes, como um conceito que não se deve de confundir com o conceito de sistemas de IA. Estes modelos apresentam uma generalidade significativa e uma capacidade de desempenhar com competência uma vasta gama de funções distintas, sendo treinados com grandes quantidades de dados, através de vários métodos (com recurso a aprendizagem supervisionada, não supervisionada ou por reforço) ⁴³⁸. Normalmente são componentes essenciais integrados em sistemas de IA, mas não constituem, por si só, sistemas de IA (sistema de IA de finalidade geral é um sistema de IA baseado num modelo de finalidade geral e com capacidade de servir para diversas finalidades ⁴³⁹).

Para estes modelos de IA de finalidade geral, o regulamento estabelece um regime específico de regras, implicando obrigações para os prestadores destes modelos (artigo 53.º) bem como regras para os seus mandatários (artigo 54.º), reservando ainda uma secção de

⁴³⁶ Considerando 132.

⁴³⁷ Considerando 133.

⁴³⁸ Cf. art. 3.º/ponto 63 e considerando 97.

⁴³⁹ Cf. art. 3.º/ponto 66.

obrigações dos prestadores nos casos onde os modelos de IA representem um risco sistémico (artigo 55.º)⁴⁴⁰.

Temos ainda um capítulo de medidas de apoio à inovação onde se regulam por exemplo os ambientes de testagem da regulamentação da inteligência artificial (artigo 57.º) e a testagem de sistemas de IA de elevado risco em condições reais fora do ambiente de testagem da regulamentação (artigo 60.º). Um outro capítulo dedicado à Governança, onde a nível europeu se observa a criação de um Comité Europeu para a Inteligência Artificial (artigo 65.º) e a nível nacional estipulam-se regras para as designações das Autoridades Nacionais Competentes. Há um capítulo que compreende medidas de acompanhamento e vigilância dos sistemas após a sua comercialização, partilha de informações e fiscalização do mercado. O regulamento termina com um capítulo onde constam os códigos de conduta e orientações da Comissão sobre a execução do presente regulamento (artigo 95.º e 96.º), um para a Delegação de Poderes e Procedimento de Comité (artigo 97.º e 98.º) e um capítulo dedicado às Sanções (artigo 99.º).

A entrada em vigor do Regulamento da Inteligência Artificial é um marco importante para a uniformização das regras relativas a inteligência artificial no seio da União Europeia, permitindo assim um desenvolvimento compatível com os seus valores e direitos fundamentais, para que desta forma a União se torne verdadeiramente competitiva ao nível mundial com as restantes potências líderes nesta temática. Ao mesmo tempo, o estipulado neste regulamento também assegura a todos os cidadãos europeus que, mesmo com a elevada incerteza que a inovação tecnológica apresenta, é possível confiarmos e nos sentirmos seguros com a crescente utilização de sistemas de IA no nosso dia a dia, desde que todas as obrigações e deveres de transparência sejam cumpridos.

Contudo, podemos afirmar que este é um documento que se apresenta incompleto, principalmente quando nos direcionamos para a temática deste projeto: a responsabilidade. É verdade que este regulamento é essencial para uma abordagem correta e uniforme a estas novas tecnologias, mas não tem menções à responsabilidade por danos causados por sistemas de inteligência artificial, o que nos levará às propostas seguintes que se focam na solução ao problema da responsabilidade⁴⁴¹. Ainda assim, é um regulamento com bastante

⁴⁴⁰ Risco sistémico é um risco específico das capacidades de elevado impacto dos modelos de IA de finalidade geral que tem um impacto significativo no mercado da União devido ao seu alcance ou devido a efeitos negativos reais ou razoavelmente previsíveis na saúde pública, segurança, na segurança pública, nos direitos fundamentais ou na sociedade no seu conjunto. Cf. art. 3.º/ponto 65.

⁴⁴¹ Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, p.256.

importância, pois com a fixação de várias obrigações e deveres e proibição de certos comportamentos, “*as normas podem funcionar como disposições legais de proteção de interesses alheios, a permitir desvelar a ilicitude delitual, sempre que não haja violação de direito absolutos, e a facilitar a prova da culpa.*”⁴⁴².

4.4. Diretiva de Responsabilidade de IA

Chegamos aqui após um longo caminho percorrido de esforços legislativos que contribuíram para a apresentação de propostas que se mostrassem mais completas e preparadas a enfrentar as várias problemáticas levantadas a propósito dos danos causados por sistemas de IA. É neste sentido que nos deparamos com as potenciais vias de solução a este problema.

Atualmente, temos duas propostas que visam sanar as dúvidas quanto à responsabilidade civil por danos causados por sistemas de inteligência artificial, tendo sido uma assinada, recentemente, pelo Presidente do Parlamento Europeu e pelo Presidente do Conselho, e estando outra ainda em discussão, são elas a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial (Diretiva de Responsabilidade de IA)⁴⁴³, e a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à responsabilidade decorrente dos produtos defeituosos⁴⁴⁴. A segunda proposta, recentemente assinada⁴⁴⁵, consiste numa revisão da Diretiva 85/374/CEE, de maneira que este regime legal se torne mais adequado à era digital e inteligência artificial, colmatando muitas das dúvidas e dificuldades que foram supramencionadas.

Assim, a potencial solução ao nosso problema passará pelo uso complementar destas duas diretivas, uma que irá rever, alterar e modernizar a diretiva dos produtos defeituosos, que, como vimos, estabelece um regime de responsabilidade objetiva, e outra focada nas adaptações às regras de responsabilidade civil extracontratual, baseada na culpa. Vejamos o que está a ser proposto.

⁴⁴² Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.362.

⁴⁴³ COM(2022) 496 final, proposta pela Comissão Europeia a 29/09/2022.

⁴⁴⁴ COM(2022) 495 final, proposta pela Comissão Europeia a 29/09/2022.

⁴⁴⁵ Assinada pelo presidente do Parlamento Europeu e pelo presidente do Conselho a 23/10/2024.

4.4.1. Proposta de Revisão da Diretiva de Produtos Defeituosos

Iniciando primeiro com a proposta que pretende rever e atualizar a Diretiva 85/374/CEE, observámos logo no seu artigo 1.º que são estabelecidas regras comuns para a responsabilidade civil dos operadores económicos pelos danos sofridos por pessoas singulares e causados por produtos defeituosos, bem como regras para a compensação desse prejuízo. Há ainda uma menção ao objetivo da diretiva que passa por contribuir para o correto funcionamento do mercado interno, enquanto assegura um nível elevado de proteção dos consumidores e restantes pessoas singulares.

Outro aspeto que esta proposta almeja é o elevado nível de harmonização. De acordo com o artigo 3.º, os Estados-Membros não devem de manter ou introduzir, na legislação nacional, disposições que se afastem das estipuladas nesta proposta de Diretiva, incluindo disposições mais ou menos rigorosas para atingir níveis diferentes de proteção para consumidores e pessoas singulares, ou seja, pretende-se uma Diretiva de harmonização máxima para se atingir um uniforme e elevado nível de proteção ⁴⁴⁶ (ponto que o Comité Económico e Social Europeu fez questão de realçar ⁴⁴⁷).

Um dos focos mais importantes desta proposta dirige-se à atualização dos pressupostos e referidas definições, assim, e tendo em conta as críticas realizadas anteriormente, encontramos positivas mudanças ao conceito de “produto” alvo. Este regime de responsabilidade objetiva deverá ser aplicado a todos os bens móveis, ainda que incorporados ou interligados com outros bens móveis ou imóveis, considerando-se como produtos, a eletricidade, os ficheiros de fabrico digital ⁴⁴⁸, as matérias-primas e o *software* (acabando com quaisquer dúvidas quanto à sua inclusão ou não no conceito de produto) ⁴⁴⁹.

⁴⁴⁶ Considerando 8.

⁴⁴⁷ Parecer do Comité Económico e Social Europeu sobre a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à responsabilidade decorrente dos produtos defeituosos de 25/01/2023 (EESC 2022-04922), ponto 2.6.

⁴⁴⁸ São “versões digitais ou modelos digitais de bens móveis, que contêm informação funcional necessária para produzir um elemento tangível ao permitir o controlo automatizado de máquinas ou ferramentas, tais como brocas, fresadoras, tornos ou impressoras 3D. Cf. art. 4.º/ ponto 2. Como por exemplo os ficheiros CAD (*computer-assisted-design*), usados para criar modelos 3D, armazenando informações e modelos de ferramentas de engenharia, arquitetura e outros *designs*. Incluído como produto, o que significa que, por ser defeituoso, se imprimir um bem 3D que cause danos então “*would give rise to liability under this Directive, when they are supplied or developed in the course of a commercial activity.*” (Considerando 14). Não sendo considerado *Software*, mas sim versões digitais de bens móveis, foi preciso acrescentá-los expressamente à noção de produto proposta. (Commission non-paper on digital elements of the proposal for a new Product Liability Directive (6201/23), p.4.).

⁴⁴⁹ “*Software is a product for the purposes of applying no-fault Liability*” (Considerando 12). Incluindo *Software* quando incorporados em bens móveis ou instalados em bens imóveis. (Considerando 6). Não é incluída uma noção de *software* nesta proposta por receio que esta não fosse à prova do futuro (“*future proof*”), contudo foram facultados alguns exemplos para entendermos do que se trata: “*Software, such as operating*

Também é proposta a definição de “componente” como qualquer elemento, tangível ou intangível, matérias-primas, ou qualquer serviço conexo incorporado ou interligado com um produto. Esta definição é incluída para assegurar que o fabricante do produto geral também responderá por quaisquer elementos incorporados sob o seu controlo, o que pode ser particularmente relevante para produtos digitais onde vários dos seus elementos podem ser introduzidos durante o processo de produção ou num estágio superveniente (através de atualizações) ⁴⁵⁰ ⁴⁵¹. Quanto aos “serviços conexos”, mencionado na definição de componente, ou seja, os serviços digitais incorporados num produto ou interligados com o mesmo, de tal modo que a sua ausência impediria que o produto desempenhasse uma ou mais das suas funções (artigo 4.º/ ponto 4), é pertinente referir que, mesmo que esta proposta de Diretiva não se aplique aos serviços, é preciso alargar o âmbito deste regime de responsabilidade objetiva aos serviços digitais (desde que estejam sob o controlo do fabricante) pois também eles podem determinar a segurança de um produto, da mesma maneira que outros componentes físicos ou digitais ⁴⁵².

No que diz respeito ao produtor, também é proposta a introdução de um novo conceito, o de “operador económico”, que segundo o ponto 16 do artigo 4.º, engloba o fabricante de um produto ou componente ⁴⁵³, o prestador de um serviço conexo, o mandatário

systems, firmware, computer programs, applications or AI systems”. (Commission non-paper on digital elements of the proposal for a new Product Liability Directive (6201/23), p.3.).

⁴⁵⁰ Commission non-paper on digital elements of the proposal for a new Product Liability Directive (6201/23), p.4.)

⁴⁵¹ Serviços conexos e outros componentes, incluindo atualizações de *software* ou *updates* devem de ser considerados sob controlo do fabricante quando este os integrou, forneceu, ou consentiu/autorizou o seu fornecimento por terceiros (no caso de numa *smart home* onde as atualizações de certos eletrodomésticos inteligentes são realizadas por terceiros, tudo consentido pelo fabricante) (Cf. art. 4.º/ ponto 5 – a) e b)). Depois de um produto ser colocado no mercado, deve de se considerar que o fabricante o tem sob o seu controlo, enquanto retiver a capacidade técnica para fornecer atualizações ou melhorias, ele próprio ou por via de terceiros. (Considerandos 15).

⁴⁵² Como por exemplo os serviços que fornecem continuamente informação rodoviária para sistemas de navegação; um serviço de monitorização de saúde dependente dos sensores de um produto físico para manter métricas da condição do seu utilizador, um serviço de controlo de temperatura que monitoriza e regula a temperatura de um frigorífico inteligente ou serviços de assistente de voz que permitem controlar um ou mais produtos por comandos de voz (serviços de acesso à internet ficam de fora destes serviços conexos por não pertencerem ao núcleo de controlo do fabricante). “*Nevertheless, a product that relies on such services and that fails to maintain safety in the event of a loss of connectivity could be found to be defective under this Directive*” (Considerando 15).

⁴⁵³ O fabricante será qualquer pessoa singular ou coletiva que: a) desenvolva, fabrique ou produza um produto; b) mande projetar ou fabricar um produto, ou quem comercialize esse produto com o seu nome e a sua marca ou através de outras características se apresente como seu fabricante; e por fim c) quem fabrique, desenvolva ou produza um produto para uso próprio. Cf. art. 4.º/ ponto 11.

⁴⁵⁴, o importador ⁴⁵⁵, o prestador de serviços de execução ⁴⁵⁶ e o distribuidor ⁴⁵⁷. Cada um terá um papel importante na hierarquia de responsabilidades como veremos adiante.

O terceiro pressuposto, é o defeito do produto, ou qualidade defeituosa, previsto no artigo 6.º, onde também verificamos um rol de novidades. Antes de mais, constatar que o critério continua a ser o mesmo, por isso um produto será defeituoso quando não oferece a segurança com que legitimamente se pode contar. Do mesmo modo, mantém-se a análise objetiva da qualidade defeituosa, não devendo de se referir à segurança que uma determinada pessoa pode legitimamente esperar, mas sim “a segurança que o público em geral pode legitimamente esperar” ⁴⁵⁸.

É nos elementos de ponderação, que devemos de ter em conta ao avaliar a segurança ou insegurança de um produto, que se verifica um alargamento mais adequado à era digital. Adequado aos sistemas de inteligência artificial encontramos adições, como por exemplo a consideração do efeito na segurança do produto de qualquer capacidade de continuar a aprender ou adquirir novas características depois de ser colocado em funcionamento (artigo 6.º/n.º 1 – c) ⁴⁵⁹), também se deve de ter conta o efeito razoavelmente previsível no produto de outros produtos que sejam utilizados em conjunto com aquele, inclusive através de interconexão (artigo 6.º/n.º 1 – d) ⁴⁶⁰), bem como os relevantes requisitos de segurança dos produtos, incluindo requisitos de cibersegurança (artigo 6.º/n.º 1 – f) ⁴⁶¹). Para além dessas circunstâncias, temos ainda as já conhecidas anteriormente, com algumas adaptações como a apresentação e características dos produtos, incluindo rotulagem, *design*, capacidades técnicas, composição e embalagem e as instruções de instalação, utilização e manutenção; a utilização ou má utilização razoável do produto; o momento em que o produto foi colocado no mercado ou em serviço ou, caso o fabricante mantenha o controlo sobre o produto, após esse momento, o momento em que o produto deixou de estar sob o controlo do fabricante

⁴⁵⁴ Uma pessoa singular ou coletiva estabelecida na União, mandatada por escrito por um fabricante para praticar determinados atos em seu nome. Cf. art. 4.º/ ponto 12.

⁴⁵⁵ A pessoa singular ou coletiva que coloca um produto proveniente de um país terceiro no mercado da União. Cf. art. 4.º/ ponto 13.

⁴⁵⁶ É a pessoa singular ou coletiva que propõe no âmbito de uma atividade comercial, pelo menos dois dos seguintes serviços: armazenagem, embalagem, endereçamento e expedição, sem ter propriedade do produto. Cf. art. 4.º/ ponto 14.

⁴⁵⁷ A pessoa singular ou coletiva na cadeia de abastecimento que disponibiliza um produto no mercado, além do fabricante ou do importador. Cf. art. 4.º/ ponto 15.

⁴⁵⁸ Considerando 22.

⁴⁵⁹ Para que se consiga refletir a expectativa legítima de que o *software* de um produto e os algoritmos subjacentes são concebidos por forma a evitar comportamentos perigosos do produto. Considerando 23.

⁴⁶⁰ Por exemplo a interconexão que se pode verificar dentro de um sistema de *smarthome*. Considerando 23.

⁴⁶¹ Um produto que não cumpra os requisitos de cibersegurança, pode ser considerado defeituoso devido a essa vulnerabilidade pelo não cumprimento de importantes requisitos. Considerando 23 e 24.

⁴⁶². Por fim, mantém-se a ideia de que um produto não pode ser considerado defeituoso apenas por um melhor já ter sido colocado ou ser posteriormente colocado no mercado ou em serviço, incluindo atualizações ou evoluções de um produto ⁴⁶³.

Verificados os requisitos, é no artigo 8.º que encontramos os critérios que determinam a responsabilidade dos operadores económicos pelos danos causados por produtos defeituosos, estipulando-se uma hierarquia de responsabilização. O fabricante envolvido no processo de produção será chamado a responder desde que se verifique que o seu produto ou componente é defeituoso ^{464 465}. Caso o fabricante incorpore no seu produto, um componente defeituoso de outro fabricante, o lesado poderá ser indemnizado pelos danos pelo fabricante do produto ou pelo fabricante do componente, ou seja, serão responsáveis em conjunto, desde que o primeiro mantenha o controlo do produto ⁴⁶⁶. Nas situações onde o fabricante do produto ou componente esteja estabelecido fora da União, serão chamados pela seguinte ordem: o importador do produto ou componente defeituoso; o mandatário escolhido pelo fabricante; e caso os últimos dois também se encontrem estabelecidos fora da União, será chamado o prestador de serviços de execução ⁴⁶⁷. Se ainda assim não for possível identificar um dos operadores económicos acabados de referir, então poderá ser chamado a responder o distribuidor do produto, desde que este não consiga identificar nenhum operador económico ou outro distribuidor, no prazo de 1 mês após o lesado lhe ter solicitado tais informações ⁴⁶⁸.

De acordo com esta proposta, deve de ser assegurado que qualquer pessoa singular que sofre danos causados por um produto defeituoso tem direito a ser indemnizada (podendo estes pedidos de indemnização ser apresentados pelas pessoas que tenham sucedido ou subrogado no direito da pessoa lesada, bem como por outra pessoa que atue em nome de uma ou mais pessoas lesadas) ⁴⁶⁹.

⁴⁶² Isto é um aspeto a ser considerado devido à quantidade de produtos na era digital que continuam sobre o controlo do fabricante, maioritariamente na forma de atualizações. Considerando 23.

⁴⁶³ Cf. art. 6.º/ n.º 2.

⁴⁶⁴ Cf. art. 7.º/ n.º 1; e considerando 26.

⁴⁶⁵ Se qualquer pessoa singular ou coletiva alterar substancialmente um produto fora do controlo do seu fabricante e posteriormente o coloca disponível no mercado ou o coloca em serviço, essa pessoa será considerada o fabricante de tal produto nos termos da alínea do n.º 1 do artigo 7.º. (Cf. art. 7.º/ n.º 4).

⁴⁶⁶ Cf. art. 7.º/ n.º 1, último parágrafo. Na eventualidade do componente defeituoso ser incorporado num produto fora da esfera de controlo do fabricante desse produto, então o lesado só poderá ser indemnizado pelo fabricante do componente. (Considerando 26).

⁴⁶⁷ Considerando 27.

⁴⁶⁸ Cf. art. 7.º/ n.º 5 – a) e b) e considerando 27.

⁴⁶⁹ Cf. art. 5.º.

Um outro aspeto essencial da Diretiva dos produtos defeituosos que poderá vir a sofrer alterações substanciais de acordo com esta proposta, está relacionado com os encargos que recaem sobre o lesado, ou seja, a prova da qualidade defeituosa, do dano sofrido e do nexo da causalidade entre a qualidade defeituosa e o dano (artigo 9.º/n.º 1).

Em primeiro lugar, estipula-se a possibilidade de o demandado ser ordenado a divulgar os elementos de prova pertinentes de que dispõe, quando solicitados pelo demandante que tenha apresentado factos e elementos de prova suficientes para fundamentar a plausibilidade da ação de indemnização (artigo 8.º/n.º 1) ⁴⁷⁰.

Em segundo lugar, face ao ónus da prova a cargo do demandante, esta proposta decidiu apresentar uma abordagem que visa aligeirar esses encargos perante a verificação de certas condições, para assim facilitar e favorecer o lesado na sua complexa procura por compensação. Relativamente ao defeito do produto, a sua existência será presumida quando se verifique uma das seguintes circunstâncias: o demandado não cumpriu a obrigação de divulgar os elementos de prova pertinentes de que dispõe, de acordo com o artigo 8.º/n.º 1 ⁴⁷¹; o demandante estabelece que o produto não cumpre os requisitos de segurança destinados a proteger os consumidores contra o risco dos danos, obrigatório segundo o direito da União ou o direito nacional ⁴⁷²; ou, por último, o demandante consegue demonstrar que o dano foi causado por uma falha manifesta do produto durante a sua utilização normal e razoável ⁴⁷³.

Também no que diz respeito ao nexo de causalidade é proposta a sua presunção sempre que se verifique que o produto é defeituoso e que o dano causado é de uma natureza normalmente compatível com o defeito em questão ⁴⁷⁴.

Ainda se poderá presumir o defeito e o nexo causal, ou ambos, quando o tribunal considere que o demandante enfrenta dificuldades excessivas, por motivos de complexidade

⁴⁷⁰ Esta ordem de disponibilização de informação por parte do demandado deve ser proporcional e referente apenas ao necessário, para que se evite informações irrelevantes para os processos. Mais, deve de se ter em conta os interesses de ambas partes na avaliação da informação disponibilizada, de maneira que, o demandante tenha acesso a informação útil à sua pretensão, mas respeitando também os interesses do demandado e terceiros envolvidos (em particular quando esteja em causa informação confidencial e “*trade secrets*”, devendo ser apenas divulgado este tipo de informação após uma ponderada avaliação dos tribunais (Considerando 32). Cf. art. 8.º/n.º 2, 3 e 4 e considerando 31 a 32.

⁴⁷¹ Esta presunção da qualidade defeituosa do produto pode servir de incentivo ao demandado, fazendo com que este cumpra a obrigação de divulgação dos elementos de prova pertinentes, resultante do artigo 8.º. (Considerando 33).

⁴⁷² De acordo com o Regulamento (UE)2023/988 do Parlamento Europeu e do Conselho de 10 de maio de 2023, relativo à segurança geral dos produtos.

⁴⁷³ Considerando 33.

⁴⁷⁴ Quando tenha sido estabelecido que o produto é defeituoso e que o dano ocorrido é, tendo por base casos semelhantes, tipicamente causado por aquele defeito em questão, então o demandante será poupado de provar totalmente o nexo causal, dando-se a sua existência como provada. (artigo 9.º/n.º 3).

tecnológica ou científica para provar elementos ⁴⁷⁵. Para que aqui se presuma um dos dois, ou ambos os pressupostos, o demandante terá de provar que o produto contribui para o dano e a probabilidade de o produto ser defeituoso ou a probabilidade da qualidade defeituosa ser a causa do dano (alíneas a) e b) do n.º 4 do artigo 9.º) ⁴⁷⁶.

Por fim, no que diz respeito a esta proposta de revisão da Diretiva 85/374/CEE, é pertinente referir também que o artigo que enumera as causas de exclusão da responsabilidade contém algumas propostas interessantes, principalmente com a introdução do conceito de “controlo do fabricante”. Tal como referido anteriormente, excluir a responsabilidade do (agora) operador económico caso este apenas provasse que, no momento da colocação do produto em circulação, se podia razoavelmente admitir a inexistência do defeito, seria problemático devido às situações onde, por causa de atualizações ou melhorias das tecnologias digitais, o bem fosse modificado posteriormente (artigo 10.º/n.º 1 – c)). Assim, é proposto que o conceito de “controlo do fabricante” seja considerado na avaliação da alínea c), o que levou a que o n.º 2 do artigo 10.º conclua que a responsabilidade do operador económico não fosse excluída caso o defeito fosse devido a um serviço conexo, a um *software* (incluindo atualizações e evoluções (*updates* e *upgrades*)), à falha de atualização de segurança ou a uma modificação substancial do produto, desde que estejam sob o controlo do fabricante ⁴⁷⁷. Do mesmo modo, também se introduz o “controlo do fabricante” como um período temporal a se ter em conta quando falamos nos riscos de desenvolvimento, ficando o fabricante isento da responsabilidade quando o estado dos conhecimentos científicos, enquanto o produto esteve sob o seu controlo, não permitiam detetar a existência da qualidade defeituosa (artigo 10.º/n.º 1 – e)) ⁴⁷⁸.

⁴⁷⁵ Parecer do Comité Económico e Social Europeu sobre a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à responsabilidade decorrente dos produtos defeituosos de 25/01/2023 (EESC 2022-04922), ponto 5.2; Cf. art. 9.º/n.º 4; Considerando 34.

⁴⁷⁶ Esta probabilidade terá de ter em conta vários fatores como a complexidade do produto, a complexidade da tecnologia utilizada (casos de *machine learning* e *deep learning*), a complexidade das informações e dados a analisar pelo demandante e a complexidade do nexo de causalidade. Todo este processo terá de ser avaliado caso a caso pelo tribunal. A razão de ser desta presunção está intrinsecamente ligada à enorme dificuldade que um comum lesado teria em fazer provas de tais factos, pelo elevado nível de conhecimento técnico e científico exigido. (Considerando 34).

⁴⁷⁷ Permitindo assim acautelar as inúmeras situações onde o defeito surge após a colocação do produto em circulação ou em serviço, continuando ainda sob o controlo do fabricante, mantendo-se assim a sua responsabilidade. Considerando 37.

⁴⁷⁸ A determinação do estado dos conhecimentos científicos e técnicos será feita por referência ao mais elevado nível de conhecimento objetivo possível, e não ao nível do efetivo conhecimento do fabricante em causa. (Considerando 37 e 38).

Temos, portanto, uma Proposta de Diretiva que visa alargar o âmbito dos pressupostos de aplicação da Diretiva dos Produtos Defeituosos, para que melhor se consiga subsumir as tecnologias digitais inteligentes no diploma, obtendo-se, conseqüentemente, uma melhor e mais justa proteção do consumidor. De igual modo, também se proporcionará uma maior segurança jurídica às empresas, que podem assim continuar a promover a inovação. Há um claro esforço legislativo no sentido de facilitar e simplificar a prova dos pressupostos essenciais deste diploma, que recai sobre o lesado, ao se estipular um conjunto de presunções da qualidade defeituosa, do dano sofrido e do nexo de causalidade entre a qualidade defeituosa e o dano. No nosso entendimento, esta abordagem que visa simplificar um possível difícil encargo, parece a mais correta na ótica da compensação da parte mais vulnerável.

4.4.2. Proposta de Diretiva de Responsabilidade de IA

Por sua vez, temos a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial, mais conhecida como Diretiva da Responsabilidade da IA (*Artificial Intelligence Liability Directive*).

O objetivo desta proposta é estipular um conjunto de regras uniformes aplicáveis a ações de indemnização de direito civil relativas a responsabilidade culposa extracontratual, configurando disposições sobre a divulgação de elementos de prova sobre sistemas de inteligência artificial de risco elevado (permitindo aos demandantes uma fundamentação para as respetivas ações), e sobre os ónus da prova (nas ações intentadas nos tribunais nacionais por danos causados por sistemas de IA) ⁴⁷⁹. De certa forma, pretende-se que as medidas aqui previstas se insiram nos sistemas de responsabilidade civil existentes em cada Estado-Membro, sem com eles entrarem em conflito (“*refletem uma abordagem que não afeta a definição de conceitos fundamentais como a “culpa” ou “dano”*”). Por outras palavras, esta proposta pretende adaptar o tradicional regime de responsabilidade extracontratual culposa, para que se atenuem ou elimine as dificuldades sentidas face aos pressupostos da culpa, dano e nexo de causalidade, criadas pelas características específicas dos sistemas de IA (como a opacidade, o comportamento autónomo e a complexidade técnica) ⁴⁸⁰. Só com a alteração ou modificação de certos aspetos, se poderá garantir que o

⁴⁷⁹ Cf. art. 1.º/n.º 1.

⁴⁸⁰ Características que tal como já mencionamos anteriormente, tornam difícil ou quase impossível ao lesado cumprir o ónus da prova a que está tradicionalmente adstrito. Considerando 3.

nível de reparação, oferecido pelas regras nacionais, no caso de danos causados por sistemas de inteligência artificial, não será inferior aos dos casos onde sistemas de IA não estão envolvidos ⁴⁸¹.

Contrariamente ao verificado a propósito da Proposta que pretende rever a Diretiva dos produtos defeituosos, aqui encontramos uma harmonização mínima, podendo os Estados-Membros adotar ou manter regras nacionais que sejam mais favoráveis para os demandantes na fundamentação de uma ação de indemnização, desde que compatível com o direito da União ⁴⁸².

No artigo 2.º da proposta, temos uma enumeração de alguns conceitos relevantes para efeitos da diretiva, constatando-se que, muitas das definições que aqui encontramos, estão já consolidadas no Regulamento de Inteligência Artificial, como por exemplo as definições de Sistema de IA (artigo 3.º/n.º 1 do RIA) e Sistema de IA de elevado risco (artigo 6.º do RIA). São também utilizados o conceito de “fornecedor” e de “utilizador” que correspondem, respetivamente, aos atuais “prestador” (artigo 3.º/n.º 3 do RIA) e “responsável pela implantação” (artigo 3.º/n.º 4 do RIA) na versão adotada do Regulamento de Inteligência Artificial. Outro aspeto de particular interesse deve-se ao dever de diligência definido como “*uma norma de conduta obrigatória, estabelecida pelo direito nacional ou da União, a fim de evitar danos aos interesses jurídicos reconhecidos a nível nacional ou da União, incluindo a vida, a integridade física, a propriedade e a proteção dos direitos fundamentais*”, o que nos permite mais uma vez fazer a ponte com as obrigações estipuladas para o prestador e restantes sujeitos mencionados no Regulamento ⁴⁸³.

Posto isto, um dos primeiros aspetos essenciais propostos é referente à divulgação dos elementos de prova, previsto no artigo 3.º. A este respeito, é estabelecido que os tribunais podem ordenar a divulgação de elementos de prova relevantes sobre sistemas de IA de risco elevado, suspeitos de terem causados danos, a quem deles disponha, a fim de permitir que as pessoas lesadas consigam determinar os fundamentos para uma ação de indemnização, face ao grande número de pessoas que estão implicadas na conceção, desenvolvimento, implantação e exploração desse tipo de sistemas. Quer isto dizer que, este novo mecanismo de divulgação de elementos de prova está limitado apenas aos sistemas de IA de elevado risco ⁴⁸⁴, que como supramencionado, implicam um rol longo de obrigações (documentação,

⁴⁸¹ Considerando 4 e 5.

⁴⁸² Cf. art. 1.º/n.º 4, e considerando 14.

⁴⁸³ Considerando 15.

⁴⁸⁴ European Data Protection Supervisor, Opinion 42/2023 on the Proposals for two Directives on AI Liability rules, outubro de 2023, p.10, ponto 23.

conservação de registos, informações dos prestadores implicados na criação e desenvolvimento do sistema de IA)⁴⁸⁵. Esta divulgação dos elementos pode ocorrer em dois cenários: primeiro numa situação pré ação, onde só haverá ordem de divulgação se o potencial demandante apresentar os factos e informações suficientes para fundamentar a plausibilidade de uma ação, e tiver antes solicitado esses elementos a quem deles disponha, sem os obter com sucesso (artigo 3.º/n.º 1, 2º paragrafo); o segundo cenário é já no decorrer de uma ação indemnizatória, onde só poderá ser ordenada a divulgação se o demandante tiver frustrado todas as tentativas proporcionadas para recolher os elementos de prova junto do demandante (artigo 3.º/n.º 2)⁴⁸⁶. Em ambas as situações, os tribunais devem de limitar a divulgação desses elementos ao que é considerado necessário e proporcionado para fundamentar uma potencial ação de indemnização⁴⁸⁷.

Na eventualidade de incumprimento da ordenação de divulgação dos elementos de prova, o n.º 5 do artigo 3.º estipula uma presunção ilidível, onde o tribunal presume que o demandado não cumpriu um dever de diligência⁴⁸⁸ pertinente no sentido da divulgação ou conservação dos elementos de prova, o que por outras palavras significa uma presunção de culpa⁴⁸⁹.

Esta proposta também apresenta uma presunção ilidível de umnexo de causalidade em caso de culpa, de acordo com o estipulado no artigo 4.º (artigo 4.º/n.º 1 e n.º 7). Para que os tribunais nacionais possam presumir o nexode causalidade entre o facto culposo do demandado e o resultado produzido por um sistema de IA ou incapacidade do sistema de produzir um resultado, tem de se verificar, cumulativamente, a culpa do demandado (provada pelo demandante ou presumida pelo tribunal nos termos do artigo 3.º/n.º 5), a probabilidade razoável deste facto culposo influenciar o resultado produzido pelo sistema de IA ou a sua incapacidade de produzir um resultado⁴⁹⁰ e por fim, o demandante tem de demonstrar que o dano originou do resultado produzido pelo sistema ou da incapacidade do sistema de produzir um resultado⁴⁹¹. Contudo, esta presunção não será aplicada quando, em ações de indemnização relativas a sistemas de IA de risco elevado, o demandado demonstre que estão razoavelmente acessíveis ao demandante elementos de prova e

⁴⁸⁵ Considerando 19 e 18.

⁴⁸⁶ Os tribunais podem também ordenar a conservação dos elementos de prova referidos, de acordo com o artigo 3.º/n.º 3.

⁴⁸⁷ Cf. art. 3.º/n.º 4.

⁴⁸⁸ Nos termos definidos no artigo 2.º/n.º 9.

⁴⁸⁹ É uma presunção que visa não só promover a divulgação dos elementos de prova como também proporcionar uma maior celeridade e eficácia nos processos judiciais. Considerando 21.

⁴⁹⁰ Cf. art. 4.º/n.º 1 – b).

⁴⁹¹ Cf. art. 4.º/n.º 1 – c).

conhecimentos especializados suficientes para provar o nexo de causalidade (artigo 4.º/n.º 4).

O artigo 4.º vai ainda mais longe, estipulando condições especiais para a verificação da presunção referida no n.º 1, quando estejam em causa demandados específicos. Nas ações de indemnização intentadas contra o prestador ou responsável pela implantação de um sistema de IA de elevado risco, só se verificará a culpa provada ou presumida do demandado quando o autor da ação tiver demonstrado que o sujeito em causa não cumpriu com os requisitos (presentes na secção 2 do capítulo 3 do RIA) e obrigações (presentes na secção 3 do capítulo 3 do RIA) a que estavam adstritos nos termos do Regulamento da Inteligência Artificial.

Mais uma vez, esta presunção ilidível, agora do nexo causal, apenas se aplicaria aos sistemas de IA de risco elevado, contudo, de acordo com o n.º 5 do artigo 4.º, é possível aplicar-se esta disposição também a sistemas não considerados de elevado risco se o tribunal considerar que é excessivamente difícil para o demandante provar o nexo de causalidade, ou seja, a utilização desta presunção em relação aos restantes sistemas que não são de elevado risco está dependente da discricionariedade dos tribunais. Este aspeto poderá ser problemático, correndo-se o risco de se aplicar a casos análogos, situações desiguais, o que colocaria em causa a harmonização pretendida neste setor, podendo mesmo ocorrer possíveis discriminações ^{492 493}.

Apesar do estado embrionário desta diretiva, que ainda virá a ser alvo de várias discussões e, conseqüentemente, alterações, já é possível enaltecer-se alguns pontos pela negativa. Neste sentido, Mafalda Miranda Barbosa refere que a solução apresentada nesta proposta tem um “*enfoque probatório*”, acabando por não sanar as dúvidas colocadas pela inteligência artificial quanto ao pressuposto do nexo de causalidade ⁴⁹⁴.

As novas regras que são aqui propostas, compartilham o sentimento de facilitar o acesso às vias de reparação para as vítimas, também presente na Proposta anteriormente mencionada, claramente plasmado nas possíveis presunções. Tal como foi dito ao longo desta dissertação, o sucesso da pretensão indemnizatória do lesado, estará dependente da verificação dos pressupostos da responsabilidade. Estando no regime da responsabilidade

⁴⁹² Permitir que os tribunais determinem por si estes conceitos indeterminados, pode levar a diversas interpretações não só ao nível nacional, como também ao nível da União.

⁴⁹³ Cristina Aragão Seia, “Artificial Intelligence: Civil Liability 3.0”, em Federica Cristiani e Cristina Elena Popa Tache, “*Tempore Mutationis in International and Comparative Law*”, Contributions to the 3rd Conference on Comparative and International Law, Bucharest, junho, 2023, p.253.

⁴⁹⁴ Mafalda Miranda Barbosa, “Ainda o futuro da responsabilidade civil pelos danos causados por sistemas de IA”, p.365.

subjetiva, a culpa apresenta-se como um dos maiores obstáculos na procura do lesado por reparação, juntamente com o nexo de causalidade. Neste sentido, ao se prever condições especiais que resultam numa presunção da culpa ou numa presunção do nexo de causalidade, facilita-se a dificuldade, encontrada pela vítima, em demonstrar pormenorizadamente de que modo é que dano foi causado por determinado facto culposo, o que “*pode ser particularmente difícil nos casos de sistemas de IA complexos*”⁴⁹⁵.

Por outras palavras, a solução proposta foca-se na criação de um regime similar ao regime do artigo 493.º, baseado numa inversão do ónus da prova que beneficia o lesado, especificamente adequado para sistemas de Inteligência Artificial. Continuamos a defender a ideia de que um regime de responsabilidade que prescindia do juízo da culpa, é o que melhor acautela estas situações, contudo, terá de haver uma manifesta alteração legislativa no sentido de englobar expressamente as tecnologias digitais. Ainda assim, o regime subjetivo, poderá ser usado complementarmente, se forem facultados expedientes, que auxiliem a difícil tarefa do lesado, como os regimes de presunções.

⁴⁹⁵ Comunicado de Imprensa da Comissão Europeia de 28 de setembro de 2022

5. Conclusões Finais

Percorrido todo este caminho, várias são as conclusões às quais podemos chegar, sendo a primeira e a mais óbvia de todas, o facto de o Direito estar constantemente a deparar-se com uma multiplicidade de novas e inesperadas situações que, suscitando diversas problemáticas, interpelam a ciência jurídica a encontrar apropriadas soluções. Isto é uma verdadeira demonstração do Direito enquanto arte voltada para o presente e para o futuro, acompanhando as variadas mutações da nossa sociedade.

Tal como seria de esperar, chegou a vez da temática da Inteligência Artificial, assunto que no nosso entendimento não chega nem muito cedo nem muito tarde, quando se antevê os saltos tecnológicos que se poderão observar nos próximos pares de ano. O foco das mentes jurídicas recai agora, em grande peso, sobre estes poderosos *softwares* compostos por algoritmos com grandes capacidades de observar o seu ambiente, recolher e interpretar dados e agir de forma inteligente, independente e imprevisível, tendo como ponto de partida experiências e interações passadas.

É precisamente face a estas novas tecnologias que se levantou nesta dissertação, no âmbito da responsabilidade civil, a questão de quem responderia por eventuais danos causados por sistemas de inteligência artificial. Ou seja, se nos depararmos com danos provocados por ações autónomas e imprevisíveis de algoritmos inteligentes, a quem é que devemos de imputar a responsabilidade e consequente obrigação de indemnizar o lesado. Neste sentido, podemos questionar se a própria máquina dotada de um algoritmo inteligente, pode responder por ela mesma, ou se teremos de encontrar um outro responsável na cadeia de criação ou distribuição destes sistemas, como o produtor, o distribuidor ou o fornecedor ou até mesmo o utilizador ou alguém que detenha um dever de vigilância e manutenção sobre esses sistemas.

Antes de mergulharmos no imenso oceano jurídico com que nos deparamos, tiramos do nosso caminho uma possível via de solução ao problema, a questão relacionada com a atribuição da personalidade jurídica a estes entes eletrónicos (questão esta que, por si só, tem conteúdo para uma complexa e longa exposição). Embora já se encontrem alguns posicionamentos na doutrina que defendem a atribuição de um estatuto de pessoas eletrónicas aos sistemas autónomos mais sofisticados, aspeto que também chegou a ser suscitado pelo próprio Parlamento Europeu em 2017, ainda nos parece tratar-se de uma possibilidade precoce e um pouco crua, quando olhamos de forma mais rigorosa para o pleno

significado de personalidade jurídica. Deste modo, o posicionamento maioritário, que também perfilhamos, tem sido a de recusar, no momento, a atribuição de tal estatuto a sistemas de inteligência artificial, pese embora, seja uma posição que cairá, eventualmente, em desuso. Somos assim afastados, por agora, da possibilidade de à própria máquina, imputarmos os danos por si causados, levando-nos a procurar uma solução um pouco mais tradicional.

Na abordagem da problemática estudada nesta dissertação, e para respondermos às questões de responsabilidade civil que nela se colocam, distinguimos e seguimos o caminho da responsabilidade civil extracontratual ou aquiliana, como o modelo utilizado quando está em causa a violação de direitos absolutos ou qualquer outra disposição legal destinada a proteger interesses alheios completamente desligados de um qualquer vínculo jurídico prévio, em detrimento da responsabilidade civil contratual que pressupõem uma relação intersubjetiva prévia entre as partes.

É assim que chegamos à regra da responsabilidade civil extracontratual subjetiva, segundo o artigo 483.º do Código Civil, que nos apresenta os cinco requisitos de verificação obrigatória e cumulativa: o facto voluntário (ou omissão), a ilicitude, a culpa, o dano e por fim o nexo de causalidade entre o facto ilícito e culposo e os danos verificados. Contudo, esta primeira tentativa de aplicação da responsabilidade civil subjetiva, peca pela sua insuficiência e inadequação face aos sistemas de inteligência artificial. Esta inadequação resulta das características intrínsecas próprias destas novas tecnologias, como a sua autonomia, a complexidade, a opacidade dos processos de tomada de decisão e a dependência de dados que dificulta a verificação dos requisitos essenciais para o recurso ao instituto da responsabilidade civil subjetiva. A imprevisibilidade (resultado dos diversos meios de aprendizagem automática e profunda), juntamente com um elevado grau de autonomia e complexidade (visto tratar-se de uma área que exige tremendos conhecimentos científicos e técnicos) associado ainda a uma multiplicidade de atores económicos envolvidos na cadeia de abastecimento, constitui um problema na procura dos responsáveis pelo dano, bem como no estabelecimento de um nexo causal adequado, o que faz com que a pretensão do lesado tenha pouca, ou quase nenhuma, probabilidade de sucesso.

Perante o perigo de, com grande probabilidade, não se conseguir indemnizar os prejuízos sofridos pelo lesado através da via subjetiva, fomos levados a olhar com bons olhos para um regime de responsabilidade civil objetiva, como caminho alternativo, pois ao prescindir do pressuposto da culpa, há, de certa forma, uma redução dos obstáculos

colocados pelas características especiais supramencionadas. Contudo, este é um regime excepcional, apenas permitido nos casos especificados na lei, o que nos obriga ao recurso aos regimes de responsabilidade objetiva já existentes.

Em primeiro lugar, equacionamos o recurso ao regime da responsabilidade objetiva do Comitente, presente no artigo 500.º do Código Civil, onde, através de uma interpretação extensiva, consideramos os sistemas de inteligência artificial como o comissário na relação de comissão. Contudo, ainda que se consiga, de uma maneira ou de outra, verificar dois dos três requisitos exigidos pelo artigo 500.º, a aplicação deste regime, à nossa problemática em questão, fica impossibilitada pela não verificação do requisito final, que implicava que estes sistemas fossem alvo de juízos ético-jurídicos de censura pelas suas condutas, aspeto que já desconsideramos ao recusar a atribuição de personalidade jurídica a estes entes tecnológicos. Somos, assim, afastados da aplicação deste regime.

Por sua vez, uma via de responsabilidade civil objetiva mais plausível para os sistemas de inteligência artificial, reside no Decreto-Lei n.º 383/89 de 6 de novembro, que estabelece o regime de responsabilidade objetiva do produtor pelos danos provocados pelos produtos defeituosos que colocou em circulação. De certa forma, nenhum outro regime fica tão perto de resolver a incógnita dos sistemas de inteligência artificial, como este, embora ainda contenha algumas imprecisões. Através do estudo dos requisitos presentes neste Decreto-Lei, na sua redação atual, só seria possível chegarmos a uma resposta parcial. Recorrendo a alguma interpretação, conseguimos aplicar este regime de responsabilidade objetiva do produtor aos casos de danos provocados por sistemas de IA considerados defeituosos, contudo, seriam deixadas de fora do raio de aplicação do Decreto-Lei, as situações onde os danos não são causados por qualquer defeito, sendo antes devido ao elevado nível de autonomia e imprevisibilidade. Não nos parece que, o critério utilizado no Decreto-Lei n.º 383/89, para averiguar a defeituosidade da IA, se coadune com as características destas tecnologias, principalmente a da autonomia e a da autoaprendizagem. Do mesmo modo, também não nos parece certo considerar como um defeito, a característica que funciona como uma componente e objetivo a prosseguir por esses mesmos sistemas. Tendo em consideração a persistência de uma lacuna parcial, concluímos que, apesar de conter uma base apropriada à temática em questão, este regime do Decreto-Lei n.º 383/89, requer uma adaptação e atualização dos pressupostos de aplicação, como por exemplo no seu conceito de defeito, de forma que, consiga, não só, englobar todos os tipos de sistemas inteligentes, mas também aperfeiçoe as disposições vigentes.

O Decreto-Lei n.º 383/89, de 6 de novembro, transpõe para o ordenamento jurídico interno, a Diretiva 85/374/CEE, de 25 de julho de 1985, o que significa que as alterações necessárias à melhor adequação deste regime de responsabilidade objetiva às tecnologias digitais emergentes impõem um esforço legislativo europeu. Felizmente, essas alterações chegaram em boa hora, com as novas propostas oriundas do Parlamento Europeu e do Conselho, ainda em discussão. Contudo, antes mesmo de lá chegarmos, há todo um caminho legislativo europeu a percorrer.

Foi com a adoção do Livro Branco que a União Europeia traçou claras diretrizes no sentido de fomentar o desenvolvimento da Inteligência Artificial na Europa, pretendendo em simultâneo assegurar que esse desenvolvimento seja fiável e seguro, respeitando os valores e direitos fundamentais europeus. É também neste momento, que se lançou a fundamental abordagem baseada no risco, que permite identificar e distinguir as várias aplicações de IA em aplicações de alto risco, como aqueles sistemas que implicam riscos significativos com a sua utilização, dos restantes sistemas que não comportam tamanha probabilidade. Este trabalho preparatório serviu como base ao mais recente ato legislativo europeu: o Regulamento de Inteligência Artificial. Com o fim de promover uma inteligência artificial, centrada no ser humano e de confiança, o RIA, partindo da abordagem do risco, claramente influenciada pelo Livro Branco, redigiu uma categorização de sistemas de IA com base em diferentes graus de riscos, resultando assim em sistemas de riscos inaceitáveis e por isso mesmo proibidos; sistemas de elevado risco, que estão adjacentes ao cumprimento de rigorosas e específicas obrigações; os sistemas de risco limitado, e por fim os sistemas de risco reduzido ou inexistente, sem quaisquer especiais exigências. Suportando-se nestas proibições e obrigações, os cidadãos europeus têm um contexto mais seguro para o desenvolvimento e utilização dos variados sistemas de inteligência artificial, servindo também como base às suas pretensões de indemnização.

Quanto à Proposta de Diretiva do Parlamento Europeu e do Conselho sobre a responsabilidade dos produtos defeituosos, verificamos que, as aclamações que foram realizadas ao longo dos anos, estão a ser ouvidas. É proposta uma atualização dos pressupostos necessários para a aplicação deste regime de responsabilidade objetiva dos produtos defeituosos, numa clara tentativa de alargamento e adequação de muitas das disposições já existentes à nova realidade tecnológica. Para além de uma revisão e atualização de alguns conceitos, também são propostas presunções de prova ilidíveis do defeito do produto, do dano e do nexo de causalidade, o que representa um claro

reconhecimento da dificuldade em provar estes elementos, principalmente quando estamos a falar de situações que envolvam sistemas de inteligência artificial. Esta tentativa de aligeirar os pesados encargos que recaem sobre o lesado, parece-nos a abordagem mais correta no sentido de melhorar a proteção da parte mais vulnerável.

No que diz respeito à outra Proposta de Diretiva do Parlamento Europeu e do Conselho, relativa à adaptação das regras de responsabilidade civil extracontratual à inteligência artificial, temos essencialmente um complemento não só para Proposta de Diretiva anterior, como também para o Regulamento de Inteligência Artificial, onde em concordância com as regras e exigências estabelecidas para sistemas de IA, se propõem duas presunções ilidíveis, uma para a culpa e outra para o nexo de causalidade. Mais uma vez, observamos uma preocupação no sentido de facilitar pesados encargos que, à partida recairiam sobre o lesado.

Com estas duas propostas (restando apenas uma em discussão), a conclusão a que chegamos é que, por parte da União Europeia, há claramente a vontade de harmonizar a legislação dos Estados-Membros no sentido de apresentar duas válidas soluções, ambas suportadas com presunções ilidíveis, que podem ser usadas de maneira complementar para solucionar os casos de danos resultantes por sistemas de inteligência artificial.

Por outras palavras, perante um caso de danos provocados por sistemas de inteligência artificial, o sucesso da pretensão indemnizatória do lesado pela via da responsabilidade subjetiva, vai depender da forma como este consegue abordar o complexo ónus da prova. Dadas as dificuldades sentidas neste sentido, julgamos ser mais apropriado o recurso a presunções legais que invertam esse ónus, o que se coaduna com o projeto constante na proposta de Diretiva de Responsabilidade de IA, quando propõe tanto uma presunção para a culpa, como uma presunção para o nexo de causalidade, embora esta última levante ainda algumas questões.

Ainda assim, a via da responsabilidade objetiva, através da Diretiva dos Produtos Defeituosos, principalmente com as alterações propostas, apresenta, na nossa opinião, uma proteção mais clara e favorável do lesado, mantendo o recurso a determinadas presunções. Reconhecemos, contudo, que esta temática beneficiaria de uma constante discussão e ponderação de maneira a melhor se acautelar o futuro.

6. Referências Bibliográficas

ANTUNES, Henrique Sousa, “A Responsabilidade Civil aplicável à Inteligência Artificial: Primeiras notas críticas sobre a resolução do Parlamento Europeu de 2020”, em *Revista de Direito da Responsabilidade*, ano 3, 2021, pp.

ANTUNES, Henrique Sousa, “Inteligência Artificial e Responsabilidade Civil: Enquadramento”, in *Revista de Direito da Responsabilidade*, ano 1, 2019, pp.139-154;

BARBOSA, Mafalda Miranda, “Ainda o Futuro da Responsabilidade Civil pelos danos causados sistemas de IA”, in *Revista de Direito da Responsabilidade*, ano 5, 2023, pp.337-369;

BARBOSA, Mafalda Miranda, “Inteligência Artificial, E-Persons e Direito: Desafios e Perspetivas”, in *Revista Jurídica Luso-Brasileira*, ano 3, n.º 6, 2017, pp.1475-1503;

BARBOSA, Mafalda Miranda, “Inteligência Artificial, Responsabilidade Civil e Causalidade: Breves Notas”, in *Revista de Direito da Responsabilidade*, ano 3, 2021, pp.605-625;

BARBOSA, Mafalda Miranda, “O Futuro da Responsabilidade Civil Desafiada pela Inteligência Artificial: As Dificuldades dos Modelos Tradicionais e Caminhos de Solução”, in *Revista de Direito da Responsabilidade*, ano 2, 2020, pp.280-326;

BOSTROM, Nick, “*Superintelligence: Paths, Dangers, Strategies*”, Oxford University Press, 2016;

CAMPOS, Juliana, “A Responsabilidade Civil do Produtor pelos danos causados por Robôs Inteligentes à luz do regime do Decreto-Lei n.º 383/89, de 6 de novembro”; em *Revista de Direito da Responsabilidade*, ano 1, 2019, pp.700-730;

ČERKA, Paulius, GRIGIENĖ, Jurgita, SIRBYKYTĖ, Gintarė, “Liability for damages caused by artificial intelligence”, *Computer Law & Security Review*, n.º 31, 2015, pp.376-389;

COELHO, Vera Lúcia, “Responsabilidade do produtor por produtos defeituosos: Tese de resistência ao DL n.º 383/89 de 6 de novembro, à luz da jurisprudência recente, 2 anos volvidos sobre a sua entrada em vigor”, in *RED – Revista Eletrónica do Direito*, n.º 2, junho, 2017;

COMISSÃO EUROPEIA, “*Avaliação da Diretiva 85/374/CEE do Conselho, de 25 de julho de 1985, relativa à aproximação das leis, regulamentos e disposições*

administrativas dos Estados-Membros relativas à responsabilidade por produtos defeituosos”, (SWD(2018 157 final), 2018;

COMISSÃO EUROPEIA, “*Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões – Aumentar a Confiança numa Inteligência Artificial centrada no Ser Humano*”, (COM(2019) 168 final), 2019;

COMISSÃO EUROPEIA, “*Livro Branco sobre Inteligência Artificial – Uma Abordagem Europeia virada para a Excelência e a Confiança*”, (COM(2020 65 final), 2020;

COMITÉ ECONÓMICO E SOCIAL EUROPEU, “*Parecer do Comité Económico e Social Europeu sobre a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa à responsabilidade decorrente dos produtos defeituosos*”, (EESC 2022-04922), 2023;

CORDEIRO, António Menezes, “*Tratado de Direito Civil VIII*”, 2ª edição revista e atualizada, Almedina, 2023;

COSTA, Ernesto, SIMÕES, Anabela, “*Inteligência Artificial: Fundamentos e Aplicações*”, 2ª edição revista e atualizada, Lisboa, FCA, 2008;

COSTA, Mário Júlio de Almeida, “*Direito das Obrigações*”, 12ª edição, reimpressão 2022, Coimbra, Almedina, 2017;

DOMINGUES, Pedro, “*The Quest for the Master Algorithm*”, Tedx Talks Conference, 2016, disponível na internet em https://www.youtube.com/watch?v=qIZ5PXLVZfo&ab_channel=TEDxTalks;

EIRANOVA ENCINAS, Emilio, “*Código Civil Alemán – Comentando BGB*”, Madrid, Marcial Pons, 1998;

EUROPEAN DATA PROTECTION SUPERVISOR, “*Opinion 42/2023 on the Proposal for two directives on AI Liability rules*”, 2023;

Expert Group on Liability and New Technologies, “*Liability for Artificial Intelligence and other emerging digital technologies*”, União Europeia, Luxemburgo, 2019;

FELÍCIO, Manuel, “*Responsabilidade Civil por Acidentes de Viação causados por Veículo Automatizado*”, in *Revista de Direito da Responsabilidade*, ano 1, 2019, pp.493-525;

FERREIRA, Ana Elisabete, “*Responsabilidade Civil Extracontratual por Danos Causados por robôs autónomos: Breves Reflexões*”, in *Revista Portuguesa do Dano Corporal*, ano XXV, n.º 17, dezembro, 2016, pp.39-64;

FORD, Martin, “*Architects of Intelligence: The truth about AI from the people building it*”, Packt Publishing, 2018;

FORD, Martin, “Rule of the Robots: “How Artificial Intelligence will Transform Everything”, SANTOS, Luís (Tradução), “*O Futuro da Inteligência Artificial: Uma transformação que já está a acontecer*”, Bertrand Editora, 2022;

GARDNER, Howard, “*Inteligências Múltiplas: A Teoria da na Prática*”, VERÍSSIMO, Maria Adriana (tradução), Veronese, Porto Alegre, Artmed Editora, 1995;

GOLUNOVA, Valentina, “Artificial Intelligence and the Right to Liberty and Security”, in *Artificial Intelligence and Human Rights*, QUINTAVALLA, Alberto, TEMPERMAN, Jeroen, Oxford University Press, 2021;

GONZÉLEZ, José Alberto, “*Responsabilidade Civil*”, 3º edição, Quid Iuris, 2013;

GONZÁLEZ, José Alberto, “Responsabilidade por danos e Inteligência Artificial (IA)”, in *Revista de Direito Comercial*, edição de 2020, pp.69-112;

GOODFELLOW, Ian, BENGIO, Yoshua, COURVILLE, Aaron, “*Deep Learning*”, The MIT Press, Cambridge, 2016;

GPAN (Grupo de Peritos de Alto Nível sobre Inteligência Artificial), “*Orientações Éticas para uma IA de Confiança*”, Comissão Europeia, abril, 2022;

GPAN (Grupo de Peritos de Alto Nível sobre Inteligência Artificial), “*Uma definição de IA: Principais Capacidades e disciplinas científicas*”, Comissão Europeia, 2022, disponível na internet em <https://digital-strategy.ec.europa.eu/pt/policies/expert-group-ai>;

HAIGH, Thomas, “Historical Reflections – Conjoined Twins: Artificial Intelligence and the Invention of Computer Science”, in *Communications of the ACM*, VOL.66, n.º 6, junho, 2023, pp.33-37;

HODGES, Andrew, “Alan Turing and the Turing Machine”, in HERKEN, Rolf, 1995, *The Universal Turing Machine: A Half-Century Survey*, 2nd edition, Nova Iorque, Springer-Verlag, 1995;

HUTCHINS, Josh, “ALPAC: The (In)famous report”, in *MT News International*, n.º 14, julho, 1996;

Joint Research Center Technical Reports, “*AI Watch: Defining Artificial Intelligence*”, Luxemburgo, 2020;

KONIG, Pascal D., et al., “Essence of AI: What is AI?”, in DI MATTEO, Larry A., PONCIBÒ, Cristina, CANNARSA, Michel, “*The Cambridge Handbook of Artificial*

Intelligence: Global Perspectives on Law and Ethics”, Cambridge University Press, 2022, pp.18-34;

KREUTZER, Ralf T., SIRRENBORG, Marie, “*Understanding Artificial Intelligence: Fundamentals, Use Cases and Methods for a Corporate AI Journey*”, Berlim, Springer, 2020;

LEITÃO, Luís Menezes, “*Direito das Obrigações: VOL.I – Introdução da Constituição das Obrigações*”, reimpressão 2023, Coimbra, Almedina, 2022;

LEYEN, Ursula von der, “*Europe`s Choice: Political Guidelines for the next European 2024-2029*”, 2024;

LIGHTHILL, James, “*Lighthill Report: Artificial Intelligence: a paper symposium*”, Science Research Council, 1973;

LIMA, Pires de, VARELA, Antunes, “*Código Civil Anotado – Volume P*”, 4º edição, reimpressão, 2010;

MAIA, Ana Rita, “A Responsabilidade Civil na Era da Inteligência Artificial – Qual o caminho?”, in *Revista Julgar Online*, maio, 2021;

MANNING, Christopher, “*Artificial Intelligence Definitions*”, Stanford University Human-Centered Artificial Intelligence, 2020, disponível na internet em <https://hai.stanford.edu/>;

MCCARTHY, John, “*History of Lisp*”, Artificial Intelligence Laboratory, Stanford University, fevereiro, 1979;

MCCARTHY, John, et al., “*A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*”, Department of Mathematics, Dartmouth College, Hanover, 1995;

MCCARTHY, John, “*What is Artificial Intelligence?*”, Computer Science Department, Stanford University, 2007, disponível na internet em <https://www-formal.stanford.edu/jmc/>;

MOREIRA, Sónia, “Considerações sobre Inteligência Artificial e Responsabilidade Civil: O Caso dos Veículos Autónomos”, in CARVALHO, Maria Miguel, “*E-Tex Yearbook – Artificial Intelligence & Robots*”, Escola de Direito da Universidade do Minho, dezembro, 2020;

MORIOKA, Masahiro, “Descartes and Artificial Intelligence”, in *Journal of Philosophy*, VOL.13, n.º 1, janeiro, 2023, pp.1-4;

NETO, Nuno Devesa, “Responsabilidade Civil pela Utilização de Robots-Advisors: A Insuficiência do atual Regime de Responsabilidade e a Necessidade de uma nova Hipótese de Responsabilidade Objetiva”, in *Revista de Direito da Responsabilidade*, ano 2, 2020, pp.909-944;

OLIVEIRA, Arlindo, “*Inteligência Artificial*”, Fundação Francisco Manuel dos Santos 2019;

PARLAMENTO EUROPEU, “Resolução do Parlamento Europeu de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica”, (2015/2103(INL));

PARLAMENTO EUROPEU, “Resolução do Parlamento Europeu de 20 de outubro de 2020, que contém recomendações à Comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial”, (2020/2014(INL)), 2020;

PINTO, Carlos Alberto Mota, “*Teoria Geral do Direito Civil*”, GESTLEGAL, 2020;

QUATTROCOLO, Serena, “*Artificial Intelligence, Computational Modelling and Criminal Proceedings: A Framework for a European Legal Discussion*”, Springer, 2020;

Research Group on the Regulation of the Digital Economy, “*Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective*”, Max Planck Institute for Innovation and Competition, Research Paper no.19-13, 2019;

RODRIGUEZ BALLELL, Teresa, “The revision of the product liability directive: a key piece in the artificial intelligence liability puzzle”, in ERA Forum, VOL.24, n.º 2, junho, 2023, pp.

RUMELHART, David E., HINTON, Geoffrey E., WILLIAMS, Ronald J., “Learning Representations by backpropagating errors”, in *Nature*, VOL.323, outubro, 1986, pp.533-536;

RUSSEL, Stuart, NORVIG, Peter, “*Artificial Intelligence: A Modern Approach*”, 4th edition, Pearson Education Limited, 2021;

SAMPAIO, Elisa Alfaia, GOMES, Paulo Jorge, “A Inteligência Artificial como auxiliar das decisões Judiciais”, In GUIMARÃES, Maria Raquel, PEDRO, Rute Teixeira, 2023, “*Direito e Inteligência Artificial*”, Almedina, 2023, pp.203-227;

SAMUEL, Arthur L., “Some Studies in Machine Learning Using the Game Checkers”, in *IBM Journal of Research and Development*, VOL.3, n.º 3, julho, 1959, pp.210-229;

SEARLE, John, “Minds, Brains and Programs”, in *Behavioral and Brains Sciences*, VOL.3, 1980, pp.417-424;

SEIA, Cristina Aragão, “Artificial Intelligence: Civil Liability 3.0”, in CRISTIANI, Federica, TACHE, Cristina Elena Popa, “*Tempore Mutationis in International and Comparative Law: Contributions to the 3rd Conference on Comparative and International Law*”, Bucareste, junho, 2023, pp.239-257;

SILVA, João Calvão da, “Responsabilidade *Civil do Produtor*”, Coimbra, Almedina, 1990;

SILVA, Nuno Sousa e, “Direito e Robótica: uma primeira aproximação”, in *Revista da Ordem dos Advogados*, ano 77, VOL.I/II, janeiro-junho, 2017, pp.487-553;

SILVA, Nuno Sousa e, “Inteligência Artificial, robots e responsabilidade civil: O que é que é diferente?”, in *Revista de Direito Civil*, ano IV, n.4, Almedina, 2019, pp.691-711;

SOUSA, Susana Aires de, “Breves Notas a Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União”, in “*Direito em Mudança: A Proposta de um Regulamento Europeu sobre Inteligência Artificial – Algumas Questões Jurídicas*”, Instituto Jurídico, Faculdade de Direito, Universidade de Coimbra, 2023;

SURDEN, Harry, “Artificial Intelligence and Law: An Overview”, in *Georgia State University Law Review*, VOL.3, n.º 4, 2019, pp.1306-1337;

SURDEN, Harry, “Machine Learning and Law”, in *Washington Law Review*, VOL.89, 2014, pp.87-115;

SUSSKIND, Richard E., “Artificial Intelligence, Expert Systems and Law”, in *The Denning Law Journal*, VOL. 5, n.º 1, 1990, pp. 105-116;

TECUCI, Gheorghe, “Artificial Intelligence”, in *WIREs Computational Statistics*, VOL.4, n.º 2, março-abril, 2012, pp.168-180;

THOMPSON, Steven John, “*Machine Law, Ethics and Morality in the Age of Artificial Intelligence*”, Engineering Science Reference (IGI Global), 2021;

TURING, Alan M., “Computing Machinery and Intelligence”, in *MIND: A Quarterly Review of Psychology and Philosophy*, VOL. LIX, n.º 236, 1950, pp. 433-460;

VARELA, João Antunes, “*Das Obrigações em Geral – VOL.I*”, reimpressão 2022, Coimbra, Almedina, 2017;

ZELEZNIKOW, John, “Using Artificial Intelligence to support legal decision-making: Is software a friend or a foe?”, in *Inteligência Artificial e Robótica: Desafios para o Direito no séc.XXI*, 1ª edição, GESTLEGAL, 2022, pp.21-43;

7. Webgrafia

ALPACA LAW, disponível na internet em <https://alpacalaw.com/sobre>;

APURA, disponível na internet em <https://apura.ai/>;

DICTIONARU OF COMPUTING, 4th edition, Oxford University Press, 1996;

e-Codex, disponível na internet em <https://www.e-codex.eu/projects>;

e-Leilões, disponível na internet em <https://www.e-leiloes.pt/faqs.aspx>;

e-Tribunal CITIUS, disponível na internet em <https://www.citius.mj.pt/portal/Faq.aspx>;

GOOGLE DEEPMIND, “AlphaGo”, disponível na internet em <https://deepmind.google/technologies/alphago/>;

IBM Documentation, “*Logical (Boolean) Operators*”, disponível na internet em <https://www.ibm.com/docs/en/zos/3.1.0?topic=expressions-logical-boolean-operators>;

IBM, “*Machine Learning*”, Design for AI, disponível na internet em <https://www.ibm.com/design/ai/basics/ml/>;

IBM, “*Neural Networks*”, disponível na internet em <https://www.ibm.com/topics/neural-networks>;

IBM, “*What are Convolutional Neural Networks*”, disponível na internet em <https://www.ibm.com/topics/convolutional-neural-networks>;

IBM, “*What are Recurrent Neural Networks*”, disponível na internet em <https://www.ibm.com/topics/recurrent-neural-networks>;

IBM, “*What is Artificial Intelligence (AI)?*”, disponível na internet em <https://www.ibm.com/topics/artificial-intelligence>;

GOOGLE DEEPMIND, “AlphaZero”, disponível na internet em <https://deepmind.google/discover/blog/alphazero-shedding-new-light-on-chess-shogi-and-go/>;

NEURALINK CORPORATION, disponível na internet em <https://neuralink.com/>;

PORTAL EUROPEU DE JUSTIÇA, disponível na internet em <https://e-justice.europa.eu/home?plang=pt&action=home>;

REUTERS, “*Future of Health*”, disponível na internet em <https://www.reuters.com/business/healthcare-pharmaceuticals/neuralinks-first-human-patient-able-control-mouse-through-thinking-musk-says-2024-02-20/>;

STOCKFISH, disponível na internet em <https://stockfishchess.org/about/>;

8. Jurisprudência Citada

Acórdão do Tribunal de Justiça de 30/09/2014, Processo n.º 368/04.0TCSNT.L1.S1, Relatora Maria Clara Sottomayor.

Acórdão do Tribunal de Justiça de 17/05/2017, Processo n.º 150/11.1TBOAZ.P1.S1, Relator António Piçarra;

Acórdão do Tribunal da Relação de Lisboa de 22/06/2021 Processo n.º 1694/18.6T8PDL.L1-7., Relator Jorge Capacete;

Acórdão Tribunal da Relação do Porto de 26/06/2023, Processo n.º 1483/21.0T8VNG.P1, Relator Miguel Baldaia de Morais.