
Maria do Céu Cunha Carrão

# Artificial Intelligence in Criminal Proceedings
# The admissibility of AI-generated evidence

Dissertation to obtain a Master's Degree in Law, in the specialty of Litigation and Arbitration Law

Supervisor:

Dr.iur Athina Sachoulidou

June 2022


Maria do Céu Cunha Carrão

# Artificial Intelligence in Criminal Proceedings
# The admissibility of AI-generated evidence

Dissertation to obtain a Master's Degree in Law, in the specialty of Litigation and Arbitration Law

Supervisor:

Dr.iur Athina Sachoulidou

June 2022

NOVA SCHOOL OF LAW

## Declaração antiplágio

Declaro por minha honra que o trabalho que apresento é original e que todas as minhas citações estão corretamente identificadas. Tenho consciência de que a utilização de elementos alheios não identificados constitui uma grave falta ética e disciplinar.

(Nome e assinatura do/a aluno/a)
Maria do Céu Cunha Carrão

Maria do Céu Cunha Carrão

## Anti plagiarism statement

I hereby declare that the work I present is my own work and that all my citations are correctly acknowledged. I am aware that the use of unacknowledged extraneous materials and sources constitutes a serious ethical and disciplinary offence.

(Student's name and signature)
Maria do Céu Cunha Carrão

Maria do Céu Cunha Carrão

**Acknowledgements**

I would like to express my deepest gratitude to my Mother, Francisca, and Father, Elisiário, for all love and support, for making me who I am. To my brothers, Hugo and Cristiano, for bringing me happiness and strength. To my dearest friends, for being light in the hardest moments. To Eduardo for the unconditional encouragement, and comfort meals. A special acknowledgement to my supervisor Professor Athina Sachoulidou, for her incredible availability and understanding.

**Agradecimentos**

Um especial e profundo agradecimento à minha Mãe, Francisca, e ao meu Pai, Elisiário, por todo o amor, carinho e apoio, por fazerem de mim a pessoa que sou hoje. Aos meus irmãos, Hugo e Cristiano, pela alegria e força que me transmitem. Aos meus amigos mais queridos, por serem luz nos momentos mais difíceis. Ao Eduardo, pelo encorajamento incondicional e pelas refeições reconfortantes. Deixo um especial agradecimento à minha orientadora Professora Athina Sachoulidou, pela sua incrível disponibilidade e compreensão.

*To the loves ones who departed to early to celebrate this accomplishment.*

*“They insisted that law and regulations were always going to be too late and never catch up with AI, when in fact norms are not about the speed but about the direction of innovation, for they should steer the proper development of a society (if we like where we are heading, we cannot go there quickly enough).”*

*Luciano Floridi - AI and Its New Winter: from Myths to Realities (2020)*

**Abstract**

During the last two decades Artificial Intelligence became ubiquitous in our lives. Revealing itself as a disruptive technology, it is already impacting important sectors of society, being a driver of the Fourth Industrial Revolution.

Artificial Intelligence is benefiting humanity, and promises innovative solutions to modern-life problems, nevertheless it has a twofold effect. Artificial Intelligence as systems that are capable to monitor their surrounding environment, autonomously collect and process data, learn and act, may constitute harm to fundamental rights, mainly when deployed to criminal justice.

This analysis will focus on the specificities of Artificial Intelligence systems, delving into the admissibility of AI-generated evidence in the Portuguese criminal evidentiary framework in light of the defence rights and structuring principles of Portuguese criminal procedure.

**Keywords:** Criminal Law and Artificial Intelligence; AI-generated evidence; Machine Evidence;

**Resumo**

Durante as duas últimas décadas, a Inteligência Artificial tornou-se uma presença constante nas nossas vidas. Ao impactar setores relevantes da sociedade, tem relevando o seu caráter disruptivo, sendo um dos motores impulsionadores da Quarta Revolução Industrial.

A Inteligência Artificial além dos seus presentes benefícios para a humanidade, promete soluções inovadoras para os problemas que afligem a sociedade contemporânea, porém a mesma comporta uma duplicidade de efeitos. Os sistemas de Inteligência Artificial pela sua capacidade de monitorizar o seu ambiente circundante, e autonomamente recolher, processar dados, aprender e agir, podem concretizar riscos para os direitos fundamentais, principalmente no contexto da justiça criminal.

Esta análise irá focar-se nas especificidades dos sistemas dotados de Inteligência Artificial, aprofundando a temática da admissibilidade da prova gerada por Inteligência Artificial no quadro probatório do Direito Processual Penal Português à luz dos direitos de defesa do arguido e dos seus princípios que norteadores.

**Palavras-Chave:** Inteligência Artificial e Processo Penal; Prova gerada por Inteligência Artificial;

# Index

## 1. Introduction: Artificial Intelligence as a driver of social change

Over the last two decades, Artificial Intelligence (AI) experienced a profound and astonishingly rapid development, becoming an integral part of our routine and the tipping point of the so-called Fourth Revolution.[1] This represents the current socio-economic and even cultural shift of paradigm resulting from the disruption caused by the extraordinary technological advancements, including but not limited to AI-driven tools and technologies. The increasing volume and variety of available data as well as the velocity of data exchange (Big Data)[2], the interconnectivity between devices (Internet of Things), and the merger of the physical, digital and biological worlds[3], combined with the speed[4] of the technological novelties and the breadth and depth with which they are affecting different sectors of our lives – not only the way we work, communicate and relate with each other, but also healthcare, environment and climate change, safety, economy and consumption patterns, politics, and manufacturing processes[5] - are the signs that we "are seeing the beginning a [sic] profound cultural revolution".[6]

Even though AI is not a new concept, its development upsurged in the aftermath of the advancements achieved by the Digital Revolution[7] which provided the necessary

[1] The entering in a Fourth Industrial Revolution pioneering studies have been deepened by Professor and Founder of the World Economic Forum Klaus Schwab and the philosopher Luciano Floridi. *See* FLORIDI, Luciano – "*The 4th Revolution: How the Infosphere is Reshaping Human Reality"*, 2014 and SCHAWB, Klaus – *"The Fourth Industrial Revolution",* 2017.

[2] Big Data is one mark of the Fourth Revolution. It refers to the amazingly growing amount of available data, its invisibility, variety and easy access, as well as to the velocity it is processed and storage capacity. The more the Internet and digital applications permeate our lives, the bigger becomes the digital footprint that fuels Big Data. The amount of data is so complex and large that is impossible to store and process with traditional methods. *See SACHOULIDOU, Athina – "OK Google: is (she) guilty?"* in Journal of Contemporary European Studies, 2021, p. 1. *And* BOUCHER, Philip – *"Artificial Intelligence: How does it work, why does it matter, and what can we do about it"*. Study for the Panel for the Future of Science and Technology, 2020, p. 1.

[3] See SCHAWB, Klaus – *"The Fourth Industrial Revolution",* 2017, p. 12.

[4] *Id* at 12, 13.

[5] *Id* at 15.

[6] FLORIDI, Luciano – "*The 4th Revolution: How the Infosphere is Reshaping Human Reality",* 2014, p. 7.

[7] Also known as the Third Industrial Revolution, which began in the second half of the XX century, was marked by the shift from analog to digital technology and by the development of electronic technology capable of manipulating and communicating information, such as by increased computing power and the opening of the World Wide Web. OLIVEIRA, Arlindo – "*Inteligência Artificial"*, 2019, pp. 12, 13.

technological background for AI development,[8] such as software, hardware,[9] and increased computing power – with the end of World War II, digital computer usage ceased to be limited to military and scientific research and extended to common human activities,[10] gradually each generation of computer hardware brought an increase in speed and capacity associated with a price decrease (Moore's Law).[11] The opening of the World Wide Web in 1993 is another milestone for AI development as it created a data-rich environment – since internet globalization, we have been generating and accessing increasing amounts of data. The data footprint left behind in the digital environment is essential for the decision-making process in AI systems. The conditions for computer engineering and AI research to flourish were met[12] – resulting in the development of more sophisticated operating systems and algorithms:[13] problem-solving software,[14] programming language (e.g., ELIZA),[15] expert systems,[16] and machine learning[17] designed to support decision-making.

[8] FLORIDI, Luciano – "*Should we be afraid of AI?"* in AEON Magazine, 2016. Available at: https://aeon.co/essays/true-ai-is-both-logically-possible-and-utterly-implausible

[9] See RUSSEL, Stuart; NORVIG, Peter – *"Artificial Intelligence: A Modern Approach",* 2010, pp. 13, 14.

[10] NILSSON, Nils J. – *"The Quest for Artificial Intelligence: A History of Ideas and Achievements"*, 2010, pp. 53, 54.

[11] In 1965 Gordon Moore observed the tendency that the number of transistors and integrated circuits in digital computer chips would double every two years, associated with a decrease in hardware price. It is still presently used to refer to the exponential increase of computing power. Today, one smartphone holds more computing power than all existing computers at the beginning of the second half of the XX century. RUSSEL, Stuart; NORVIG, Peter – *Artificial Intelligence: A Modern Approach*. 2010, p. 14; OLIVEIRA, Arlindo – *Inteligência Artificial,* 2019, p. 35.

[12] An Executive's guide to AI: *Why AI now?* Available at: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai

[13] "Algorithm as a set of rules defining how to perform a task or solve a problem. In the context of AI, this usually refers to computer code defining how to process data." BOUCHER, Philip – *"Artificial Intelligence: How does it work, why does it matter, and what can we do about it"*. Study for the Panel for the Future of Science and Technology, 2020, p. VI.

[14] E.g.: Logic Theorist (1957) and GPS (1961 - General Problem Solver), designed to simulate human problem-solving protocols. RUSSEL, Stuart; NORVIG, Peter – *Artificial Intelligence: A Modern Approach*, 2010, p. 18; GUGERTY, Leo – *"Newell and Simon's Logic Theorist: Historical Background and Impact on Cognitive Modeling"*, 2006 available in https://www.researchgate.net/publication/276216226_Newell_and_Simon's_Logic_Theorist_Historical_Background_and_Impact_on_Cognitive_Modeling

[15] ELIZA was one of the first language software capable to have a conversation in natural language with a human, being a predecessor of today's chatbots. OLIVEIRA, Arlindo – "*Inteligência Artificial"*, 2019, pp. 55-57.

[16] *See*: BOUCHER, Philip – *Artificial Intelligence: How does it work, why does it matter, and what can we do about it*. Study for the Panel for the Future of Science and Technology, 2020, p. 2; BOSTROM, Nick – "*Superintelligence: Paths, Dangers, Strategies"*, 2014, p. 3, 7.

[17] BOUCHER *supra note* 16 at 152.

As a repercussion of this technological progress, and given its capacity to make our lives easier AI became ubiquitous in our lives, embodied in tools and devices designed to assist humans in daily tasks and enhance productivity it has assumed the role of "the defining technology of the last decade and probably also for the next".[18] The main objective of AI technology is to increase efficiency and assist humans across a great array of daily tasks, which is possible through the increasing automation facilitated by emerging digital technologies. AI's increasing pervasiveness is illustrated on many daily occasions: when we set our GPS to calculate the shortest route to get to work; whenever we use a search browser; in the automatic correctors of our smartphones; email spam detection; automatic translation tools; digital assistants (such as Alexa, Siri, Cortana); use of biometric data to access our personal devices, recommender systems and through targeted advertisement.[19]

Broadly defined as machines that are capable to act like humans there is still no consensual definition for AI, which for the purposes of this introductory chapter, is considered an umbrella term that includes technologies and tools capable "of displaying intelligent behavior by analyzing and adapting to their environment and providing autonomous outputs with little or no human control or supervision to achieve a specific goal".[20]

Different from conventional digital technologies, AI systems present several distinctive functions; they can be: *descriptive* and *diagnostic*, as they are capable to perceive, analyze and collect the data present in the surrounding environment through sensors (cameras, microphones, keyboards) as well as sensors of physical quantities (temperature, distance, speed, force); *predictive* in the sense of forecasting possible events through reasoning and learning methods; *prescriptive* in the sense of performing *reasoning and decision-making* tasks and *taking action* in accordance with their decisions.[21] AI systems are not limited

[18] *See* COM (2020) 65 final. Brussels, 19.02.2020 – "*White Paper on Artificial Intelligence – A European approach to excellence and trust"*, p. 2. And BOUCHER, Philip – "*Artificial Intelligence: How does it work, why does it matter, and what can we do about it"* Study for the Panel for the Future of Science and Technology, 2020, p. 1.

[19] See generally: SHIN, Donghee – *" How do Users Interact with Algorithm Recommender Systems? The Interaction of Users, Algorithms, and Performance"* in Computer in Human Behaviour, vol. 109, 2020.

[20] BOUCHER *supra note* 16, at III. *And* COM (2018) 237 final. Brussels, 25.04.2018 – *Artificial Intelligence for Europe*, p. 1.

[21] GIUFFRIDA, Iria – "*Liability for AI Decision-Making: Some Legal and Ethical Considerations",* 2019, p.440 and The European Commission's High-Level Expert Group on Artificial Intelligence – "*A Definition of AI:*

to processing information from a static, already set database, they can also be creative, perform tasks autonomously and modify their surrounding environment, standing out for their adaptability and autonomy, intrinsic features enabled by their learning capacity.[22]

One popular example of how AI systems work is *Roomba*, an AI embodied cleaning robot. Its sensors will capture if there is any object on the floor that must be avoided and recognize dirt, through its reasoning modules the robot will interpret the collected data and decide if the floor should be cleaned, at the end the robot will act according to its decision by cleaning or staying put.[23] Besides detecting dirt, it adapts to the household needs and routine, by cleaning when there is less activity inside the house and focusing on places where the dirt is mostly found.

Another distinctive characteristic of AI is that it "does not perform in an informational vacuum" but in a technological combined environment.[24] AI systems are just a part of multiple technologies that interact with each other and with the kinetic world around them. It is the combination between AI, the Internet, digital devices, and robotics that creates the "AI ecosystem"[25] and expands the usefulness and commercial potential of AI.

Smart homes are a clear example of how AI and interconnected devices may contribute to a more efficient, safer, and comfortable environment. Network-connected smart home appliances include, for instance: web-connected cleaning robots that are able to automatically identify what surfaces need to be clean and follow a cleaning schedule; smart thermostats capable to adapt room temperature according to the number of persons in a room and to learn from behavior to provide a more sustainable energy use; voice

*Main Capabilities and Scientific Disciplines",* 2018, p. 2-3. *See also* LIGETI, Katalin – *"Artificial Intelligence and Criminal Justice"* in AIDP-IAPL International Congress of Penal Law, 2019, p. 2.

[22] PAGALLO, Ugo – *"Research Handbook on the Law of Artificial Intelligence",* 2018, p. XXIV. *See also* The European Commission's High-Level Expert Group on Artificial Intelligence – *A Definition of AI: Main Capabilities and Scientific Disciplines*, 2018, p. 3.

[23] SURDEN, Harry; WILLIAMS, Mary-Anne – *"Technological Opacity, Predictability, and Self-Driving Cars"* in Cardozo Law Review, Vol. 38, 2016, p. 131.

[24] See, GIUFFRIDA, Iria – *"Liability for AI Decision-Making: Some Legal and Ethical Considerations"*, 2019, p. 441 and GIUFFRIDA, Iria; LEDERER,Frederic; VERMERYS, Nicolas – "*A Legal Perspective on the Trials and Tribulations of AI: How Artificial Intelligence, the Internet of Things, Smart Contracts and Other Technologies Will Affect the Law"* in Case Western Reserve Law Review, Vol. 68, 2018, p. 760.

[25] See GIUFFRIDA, Iria –"*Liability for AI Decision-Making: Some Legal and Ethical Considerations"*, 2019, p. 442.

assistants that obey to commands; smart refrigerators capable to control groceries' storage and send a grocery list to the user's smartphone or suggest recipes with the available content; and fire and carbon monoxide detectors that, besides triggering a sound alarm, send an alert message to the user's smartphone.

Moreover, the merger of the physical, the digital, and the biological world, enabled by AI development, results in promising advances in healthcare by improving disease diagnosis and treatments. Projects such as MRI and Ultrasound Robotic Assisted Biopsy (MURAB)[26], BioMind[27], and Corti's AI triage assistant AUDIA[28] are examples of how AI facilitates early and effective disease diagnosis. The fusion between the human body and digital technology is a step closer to the development of Brain-Computer Interfaces (BCIs) which may revolutionize the treatment of neurological disorders.[29]

Road safety and mobility are also benefiting from AI as automated vehicles and their advanced safety systems, such as drowsiness warning and intelligent speed assistance,[30] which are the newest bet of the European Commission to prevent road accidents caused by human error.[31] AI can also contribute to fighting climate change, with the building of smart cities focusing on energy-efficient buildings and the optimization of renewable energy use through distributed energy grids, and smart farming involving automated data collection and corrective actions to allow early detection of crop diseases in order to avoid the use of pesticides.[32]

[26] See more at: https://www.murabproject.eu/about-murab/

[27] BioMind is an AI company that develops intelligent solutions in medical imaging. More information available at: https://www.biomind.ai/product/

[28] AUDIA is an AI assistant that guides the triaging process during emergency calls with an accuracy rate of 92%. See: https://www.corti.ai/solutions/call-center-triage

[29] The NeuraLink is a revolutionary and ambitious project funded by Elon Musk it is under research and nearing the testing phase. https://neuralink.com/applications/
See also about this subject PELERIGO, Vanessa – *"Brain Computer Interface – Uma Primeira Abordagem"* in Anatomy of Crime: Journal of Law and Crime Sciences, n.º 12, 2020, pp. 75-80.

[30] Regulation (EU) 2019/2144 of 27 November 2019, art.º 3.

[31] COM (2018) 293 final. Brussels, 17.05.2018 - *"EUROPE ON THE MOVE Sustainable Mobility for Europe: safe, connected, and clean",* p. 2.

[32] HERWEIJER, Celine – *"8 Ways AI can Help Save the Planet",* 2018 available at: https://www.weforum.org/agenda/2018/01/8-ways-ai-can-help-save-the-planet/

Notwithstanding the fact that AI is a strategic technology that offers several benefits to citizens and the society as a whole,[33] it is revealing itself as a disruptive and transforming technology. We have reached the point where technology reveals to be much more than a mere assistance tool, it has become an environmental, social, anthropological, and interpretative phenomenon.[34] In association with ICTs, AI is rapidly disrupting traditional social patterns and generating new social dynamics. The way we communicate and relate to each other, work, consume and spend our leisure time is changing. Humanity is now *hyperconnected* and living an *onlife* experience.[35] Internet connectivity ceased to be limited to computers, it now extends to smartphones, home appliances, vehicles, as well as to industry and commercial tools, this phenomenon is known as the Internet of Things[36] (IoT). AI and the IoT can combine the physical and virtual worlds, generating a new smart environment that senses, analyses, adapts, and makes decisions.[37]

As a mirror of the sociocultural context, Law is not left untouched by AI effects. The digital and technological turn not only brought innovation and efficiency to the realm of traditional juridic professions but also pervades progressively (criminal) justice administration and starts raising liability assessment-related questions or evidence assessment-related concerns – an impact that is visible, *inter alia*, in the realm of criminal law.

The embodiment of AI endowed systems in the court's decision-making process is already taking place in some countries and raising debate.[38] The growing existence of automated vehicles and autonomous robots capable of acting without human control or

[33] COM (2020) 65 final. Brussels, 19.02.2020 – *"White Paper on Artificial Intelligence – A European approach to excellence and trust"*, p. 25.
[34] FLORIDI, Luciano – *"The Fourth Revolution. How the Infosphere is reshaping humanity",* 2014, p. 7.
[35] *Hyperconnected society* and *onlife experience* are terms coined by the philosopher Luciano Floridi to represent the Internet of Things phenomenon. FLORIDI, Luciano – "*The Fourth Revolution: How the Infosphere is Reshaping Humanity"*, 2014, p. 50: "The digital-online world is spilling over into the analogue-offline world and merging with it. This recent phenomenon is variously known as 'Ubiquitous Computing', 'Ambient Intelligence', 'The Internet of Things', or 'Web-augmented things'. I prefer to refer to it as the *onlife* experience."
[36] ISOfocus September-October 2016 – ISSN 2226-1095, p. 15.
[37] *Id.*
[38] As an example refer to United States, where AI systems designed to support the judge in decision-making processes are diffused. The case Loomis v. Wiscosin started a debate regarding the use of predictive algorithms in criminal justice. This subject will be approached in chapter 3. *See generally*: ZAVRŠNIK, Aleš – *"Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings" in* European Journal of Criminology n.º I-2, 2019.

previous input is consequently generating a new social context of distributed morality and responsibility, challenging the traditional paradigm of criminal liability in cases where intelligent devices cause harm.[39]

The capability of intelligent machines to collect data and to react autonomously according to their surrounding environment may also generate the so-called machine evidence.[40] This implies that one may resort to intelligent devices as a valuable source of information in the course of criminal proceedings, giving place to a new generation of evidence. Taking into consideration the socio-cultural context in continental Europe, the focus of this thesis lies on the admissibility of AI-generated evidence - a problem that will be examined in the light of the Portuguese criminal procedural law by reference to the existing legal framework and the fundamental principles governing the criminal procedure. The following analysis will be based on the example of the advanced safety systems which will be mandatory in new cars circulating European roads, starting in July 2022.[41]

The main analysis will be divided into four parts. First, this thesis will delve into the definition of AI and the main distinctive traits of AI-based systems in terms of the leading concepts of the following analysis (Section 2). Subsequently, it will discuss how AI and criminal justice may intersect with each other and provide a definition of AI-generated evidence (Section 3). Against this background, it will examine subsequently the admissibility of AI-generated evidence in the light of the Portuguese criminal evidentiary rules (Section 4) and its compliance with criminal procedural rights (with a focus on defence rights) (Section 5). This analysis will take into consideration the EU approach to AI and the recently published Proposal for a Regulation on AI.[42]

---

[39] PAGALLO, Ugo; QUATTROCOLLO, Serena – "*The Impact of AI on Criminal Law and its Twofolds Procedures"* in Research Handbook on the Law of Artificial Intelligence, 2018, p. 386.
For this subject see also: GLESS, Sabine; SILVERMAN, Emily; WEIGEND, Thomas – "*If Robots Cause Harm, Who is to Blame? Self-Driving Cars and Criminal Liability"*, 2016. And FLORIDI, Luciano – "*Distributed Morality in an Information Society"* in Science and Engineering Ethics, n.º 19, 2012.

[40] See GLESS, Sabine – "*AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials"* in Georgetown Journal of International Law, Vol. 51, n.º 2, 2020, p. 195. And NUTTER, Patrick – *"Machine Learning Evidence: Admissibility and Weight in Journal of Constitutional Law",* vol. 21:3, 2019, p. 922. See generally: ROTH, Andrea – "*Machine Testimony*" in Yale Law Journal, Vol. 126, n.º 1, 2017.

[41] Regulation (EU) 2019/2144 of 27 November 2019, art.º 19.

[42] COM (2021) 206 final. Brussels, 21.04.2021 – Proposal for a Regulation of The European Parliament and of the Council. Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts.

## 2. Seeking to define AI: the *status quo*

As previously mentioned, AI is not a new concept, its first marks date back to 1950 when Alan Turing wondered whether machines can think in his paper entitled "Computer Machinery and Intelligence".[43] In this paper's chapter entitled "Imitation Game", Turing presented the test where it was proposed for a machine to deceive a human interrogator by passing successfully by another human.[44] If the machine indistinguishably passed this test, then it would be a "thinking machine". A few years later, John McCarthy coined the term "Artificial Intelligence" by defining it as the science and engineering of making intelligent machines.[45]

Today, there is still no consensus regarding the definition of AI either in the scientific or in the legal scholarship. From Alan Turing's imitation game to John McCarthy's and Marvin Minsky's intelligent machines, AI remains a *to be* defined concept. The interdisciplinarity and the amazingly rapid evolution of AI-based technologies are constantly moving the frontier of what AI is, so what could be considered AI a few years ago, is now far from what we consider close to being AI.

Besides its interdisciplinarity, the rapidity with which AI technology evolves prevents the existence of a stable consensual definition, giving place to the phenomenon entitled odd paradox.[46] There is, however, unanimity as regards one thing: AI resides in the emulation of human intelligence by a machine, for instance, John McCarthy considered "the ultimate effort (of AI) is to make computer programs that can solve problems and achieve goals in the world as well as humans", [47] and Marvin Minsky refers to AI as "the science

[43] TURING, Alan. M – "*Computer Machinery and Intelligence*. Mind – A Quarterly Review of Psychology and Philosophy", 1950. Available at: https://www.csee.umbc.edu/courses/471/papers/turing.pdf
[44] TURING, Alan. M – "*Computer Machinery and Intelligence*. Mind – A Quarterly Review of Psychology and Philosophy", 1950, p. 433.
[45] MCCARTHY, John – *"What is Artificial Intelligence"*, available at: whatisai.dvi (unimi.it).
[46] *See* STONE, Peter, et.al – "*Artificial Intelligence and Life in 2030 - One Hundred Year Study on Artificial Intelligence". Report of the 2015 Study Panel*, 2016, p. 12. And PELERIGO, Vanessa – "*Brain-Computer Interface – Uma Primeira Abordagem"* in Anatomy of Crime: Journal of Law and Criminal Science, nº. 12, 2020, p. 71.
[47] McCarthy, John – *"What is Artificial Intelligence",* 2007, p. 5. *See also*: SAMOILI, Sofia, et al. – "*AI Watch Defining Artificial Intelligence: Towards an operational definition and taxonomy of artificial intelligence"*. JRC Technical Reports, 2020, p. 4. See also: *And* BARTRAM, Robert, et al – "*The Age of Artificial Intelligence"*. ISO Focus Magazine. Nov-Dec 2019, p. 19 and 21.
About this subject see also: NEWELL, Allen and HERBERT, Simon – *"Computer Science as Empirical Inquiry: Symbols and Search",* 1976, p. 116: (While illustrating AI through physical symbol systems capable of

of making machines do things that would require intelligence if done by men".[48] Even the term AI itself is a metaphor for the human quality of intelligence.[49] However, intelligence itself is a complex phenomenon as a lot of the human brain and mind are still to be uncovered.[50]

Scientific scholarship concluded that to define AI, it would be necessary to oversimplify the concept of intelligence, and find a working definition of AI.[51] In the article titled "What is AI, Anyway?" Roger Schank concludes that one way to solve the lack of an AI definition is "to list some features that we would expect an intelligent entity to have"[52], each feature would be an integral part of intelligence. Communication, internal knowledge, world knowledge, intentionality, and creativity should be the critical features of an intelligent machine.[53]

A leading approach is provided by Stuart J. Russell and Peter Norvig, who adopted the *rational agent approach* by defining AI as the field of building rational computer agents that act to achieve the best outcome or the best-expected outcome in case of uncertainty.[54] This means that for a computer or machine to be rational, it implies more than merely acting, a rational agent must act correctly or adequately when confronted with a specific situation. That depends on characteristics such as perceiving the surrounding environment

general intelligence action) "By *general intelligent action* we wish to indicate the same scope of intelligence as we see in humian [sic] action"; NILSSON, Nils, J – "*The Quest for Artificial Intelligence – A History of Ideas and Achievements"*, 2012, p. 13. *And* STONE, Peter, et.al – *Artificial Intelligence and Life in 2030 - One Hundred Year Study on Artificial Intelligence. Report of the 2015 Study Panel*, 2016, p. 13.

[48] MINSKY, Marvin – "*Semantic information Processing"*, 2015, p. V.

[49] BOUCHER, Philip – *"What If We Chose New Metaphors For Artificial Intelligence?"*, 2021, p. 1.

[50] SCHERER, U Mattew – *"Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies",* 2016, p. 360.

[51] SAMOILI *supra note 51 at 4. See also* WANG, Pei – "*On the Working Definition of Intelligence"* in Center for Research on Concepts and Cognition Indiana University, 1995, p. 2: According to Pei Wang, it is obvious that after decades dedicated to studying intelligence, we still do not know much about it, therefore we must focus on finding a working definition, one that is concrete enough that we can directly work with.

[52] SCHANK, Roger C. – "*What is AI, Anyway?"* in AI Magazine, vol. 8, n.º 4, 1987, p. 60.

[53] *Id.*

[54] According to Stuart J.Russel and Peter Norvig – "*Artificial Intelligence: A Modern Approach"*, 2010, p.3 : "The quest for artificial flight succeeded when engineers and inventors stopped imitating birds and started using wind tunnels and learning about aerodynamics. Aeronautical engineering texts do not define the goal of their field as *making machines that fly so exactly like pigeons that they can fool even other pigeons*." Through this analogy the authors conclude that AI does not need to simulate to perfection human behavior and mind, AI must act humanly by *doing the right thing* when confronted with a wide variety of novel situations – rationality-, which is possible due to a combination of mathematics, engineering, and control theory.

autonomously, persisting over a prolonged time period, adapting to change, and creating and pursuing goals.[55] For instance, Nils. J. Nilsson defined AI as the activity devoted to making intelligent machines, and intelligence as the quality that enables an entity to function appropriately and with foresight in its environment, which requires many different capabilities, depending on the existing environment.[56]

This means that although rationality is a significant part of the concept of AI, it is not its only element. AI must be composed of the characteristics that allow it to achieve rationality, such as perception, reasoning, autonomy, learning, communicating, and acting in complex environments.[57]

A similar approach was also adopted by the High-Level Expert Group on AI (HLEG),[58] appointed by the European Commission (EC) to advise on the guidelines for the implementation of the European AI strategy.[59] The HLEG's definition is the starting point for the development of a definition of AI at the EU level which consists of an expanded and more technically developed version of the brief definition of AI presented in the European Commission's Communication "AI for Europe".[60] The group uses the term *AI system* (as they are usually embedded as components of larger systems) to refer to any AI-based component, software, and/or hardware, designed by humans that are able to act

[55] *Id.* at 4.
[56] NILSSON, Nils, J – *"The Quest for Artificial Intelligence – A History of Ideas and Achievements"*, 2012, p. 13.
[57] NILLS, Nilson J – *"Artificial Intelligence: A New Synthesis"*, 1998, p. 1.
[58] In 2018 the European Commission created the HLEG on AI, a group of 52 experts tasked to support and advise on the implementation of guidelines to the European AI strategy. The HLEG besides providing the Ethics Guidelines for a Trustworthy AI had also contributed to a definition of AI at the European level. *See*: The European Commission's High-Level Expert Group on Artificial Intelligence – "*A Definition of AI: Main Capabilities and Scientific Disciplines"*. 2018 and COM (2021) 206 final. Brussels, 21.04.2021 - *Proposal for a Regulation on Artificial Intelligence (Artificial Intelligence Act)*, p. 8, about the collection and use of expertise.
[59] The European AI strategy is part of the EU's main goal to create a Digital Single Market, see COM (2018) 237 final. Brussels, 25.04.2018 – *Artificial Intelligence for Europe* and COM (2015) 192 final. Brussels, 06.05.2015 – *A Digital Single Market Strategy for Europe*
[60] COM (2018) 237 final. Brussels, 25.04.2018 – *Artificial Intelligence for Europe*, p. 1: "Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)."

with rationality, by pointing out the three main capabilities that these systems must have to be considered rational: perception, reasoning/decision-making, and actuation.[61] According to the HLEG a purely rational system will not be able to always take the best action as it would lack the necessary capacity of adapting its behavior to achieve better goals, hence we must refer to *learning rational systems* when defining AI.[62]

Therefore, besides rationality, the HLEG refers to AI as a scientific discipline by identifying the main two intrinsic techniques that are currently used to build AI systems: *reasoning/decision-making techniques* which imply transforming the collected data into knowledge usable by the machine and making decisions through a combination of making inferences, planning and scheduling activities, solution search; and *machine learning techniques*, such as neural networks, deep learning, decision trees, that allow AI systems to learn how to solve multiple not specified problems, adapt and optimize their predictions and responses:[63]

"Artificial intelligence (AI) refers to systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data, and deciding the best action(s) to take (according to pre-defined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analyzing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors, and actuators, as well as the integration of all other techniques into cyber-physical systems)." [64] This definition was considered by AI Watch experts as highly technical, but very comprehensive "for incorporating all the essential aspects of AI such

[61] The European Commission's High-Level Expert Group on Artificial Intelligence – "*A Definition of AI: Main Capabilities and Scientific Disciplines"*, 2018, p. 1-3.
[62] *Id.* at 3.
[63] *Id.* at 4.
[64] The European Commission's High-Level Expert Group on Artificial Intelligence – *A Definition of AI: Main Capabilities and Scientific Disciplines*, 2018, p.7.

as perception, understanding, interpretation, and adaptive behaviour, whereas other definitions don't address them entirely".[65]

Against this background, one may conclude the following: Mere computing power does not suffice to distinguish AI *stricto sensu* from what is called brute force methods. There are many machines that –despite surpassing human thinking capacity– are not AI-driven. The historic example is "Deep Blue", the supercomputer which after beating the chess champion Kasparov, was not considered AI by the International Business Machines Corporation (IBM).[66] Its success relied mainly on its advanced calculating power, as it was able to process 200 million possible moves and determine the optimal next move looking 20 moves ahead, which was impressive but lacked the necessary AI's degree of autonomy, adaptability, and foresight.[67] Some researchers understand Deep Blue relied on expert systems,[68] which require human experts to encode their knowledge in a way computers can understand. These systems belong to the first wave of symbolic AI which places significant constraints on their degree of autonomy, as these systems can only perform tasks automatically in the ways they are instructed, and their improvement is limited to direct human intervention.[69] Hence symbolic AI is less effective for complex problems and works mainly in an "if-then-else rule", therefore expert systems are not true AI as they are constrained to very limited environments.[70] Deep Blue worked in a constrained environment based on the calculation of quantifiable possibilities, for it to be

[65] SAMOILI, Sofia, et al. – "*AI Watch Defining Artificial Intelligence: Towards an operational definition and taxonomy of artificial intelligence"*. JRC Technical Reports, 2020, p. 8.

[66] In May 1997, the chess champion Garry Kasparov was defeated by IBM's supercomputer Deep Blue. KORF, Richard E. – "*Does Deep-Blue use AI?"* in AAAI Technical Report, 1997, p. 1 : "Surprisingly, there was almost no mention of artificial intelligence in any IBM web pages. Even more surprisingly, IBM's answer to this question was "no"!" *See* also STONE, Peter, et.al – *Artificial Intelligence and Life in 2030 - One Hundred Year Study on Artificial Intelligence. Report of the 2015 Study Panel*, 2016, p. 13: "Curiously, no sooner had AI caught up with its elusive target than Deep Blue was portrayed as a collection of *brute force methods* that wasn't real intelligence. (…) Once again, the frontier had moved.

[67] NILSSON, Nils, J – "The *Quest for Artificial Intelligence – A History of Ideas and achievements*", 2012, p. 594, 595. *See also* HAENLEIN, Michael; KAPLAN, Andreas – *"A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence"*, 2019, p. 4 and LIPTON, Zachary C. – *"From AI to ML to AI: On Swirling Nomenclature & Slurried Thought"*, 2018. Available at: https://www.approximatelycorrect.com/2018/06/05/ai-ml-ai-swirling-nomenclature-slurried-thought/

[68] HAENLEIN, Michael; KAPLAN Andreas, – "*Siri, Siri, in my hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence* in Business Horizons", 2019, p. 18: "IBM's famous Deep Blue chess-playing algorithm, which beat Garry Kasparov in the late 1990s, was not AI but an expert system."

[69] BOUCHER, Philip – *Artificial Intelligence: How does it work, why does it matter, and what can we do about it"*. Study for the Panel for the Future of Science and Technology, 2016, pp. 2,3.

[70] HAENLEIN; KAPLAN, supra note 67 at 4.

considered true AI it should be able to collect and interpret data, learn from such data, use the learnings to achieve its goals and tasks through adaptation as these are the characteristics that provide AI its defining autonomy and foresight.[71]

Today's scientific knowledge and technology development are still far from achieving the "thinking machine" predicted by Alan Turing. What would be defined as "strong" AI or Artificial General Intelligence (AGI) – capable to outperform human intelligence in several generic areas and contexts, able to reason, plan and solve "out of the box problems" autonomously for tasks they were never designed – is a rather theoretical idea. And the same applies to the so-called Artificial Super Intelligence (ASI) - the peak of human brain emulation[72] by an artificial system, where machines would have equalled human intelligence implying self-awareness and consciousness, scientific creativity, social skills and general wisdom.[73] Notwithstanding the difficulties in achieving strong AI, the present technological advance and knowledge on brain structure and functionality[74] allow what we call today narrow or weak AI, which consists of intelligent systems trained and allocated to specific tasks, designed to aid humans, instead of duplicating human mental activities.[75] Thus, while today's narrow AI consists of systems designed to be deployed to specific situations to assist humans in performing certain tasks (e.g., automated driving, vacuum robots, digital assistants), strong AI that seeks to produce systems that can perform the exact same activities as humans by exhibiting aware cognition and capacity to understand its own mental states and subjective experiences is

[71] NILSSON, Nils, J – "The *Quest for Artificial Intelligence – A History of Ideas and achievements"*, 2012, p. 594: "Deep Blue, as it stands today, is not a learning system. It is therefore not capable of utilizing artificial intelligence to either learn from its opponent or *think* about the current position of the chessboard." And HAENLEIN, Michael; KAPLAN, Andreas – *A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence*, 2019, p. 4.

[72] BOUCHER, Philip – *Artificial Intelligence: How does it work, why does it matter, and what can we do about it"*. Study for the Panel for the Future of Science and Technology, 2016, p. 16. *See also* BOSTROM, Nick – *Superintelligence: Paths, Dangers, Strategies*, 2014, p. 30. *And* OLIVEIRA, Arlindo – "*Inteligência Artificial",* 2019, p. 81.

[73] European Commission for the Efficiency of Justice (CEPEJ) – *"European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment"*, 2018, appendix I, p. 31. *also* HAENLEIN, Michael; KAPLAN Andreas, – *Siri, Siri, in my hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence* in Business Horizons", 2019, p. 16.

[74] BOUCHER, Philip – Artificial Intelligence: How does it work, why does it matter, and what can we do about it. Study for the Panel for the Future of Science and Technology, 2020, p. 4.

[75] See NILSSON, Nils, J – *"The Quest for Artificial Intelligence – A History of Ideas and achievements"*, 2012, p. 388, 389 and BOUCHER, Philip – *Artificial Intelligence: How does it work, why does it matter, and what can we do about it"*. Study for the Panel for the Future of Science and Technology, 2016, p. VII

still far from existing.[76] Nonetheless, what is called "weak" AI is a pervasive technology that, as already shown (Section 1), has already revolutionized several areas of our lives.[77]

### 2.1 Autonomy and learning machines

Autonomy is a common operational element of most AI definitions, one of the corollaries of building intelligent machines capable to act like humans imply their capacity to act autonomously.[78] Although this trait is contested,[79] as autonomy is a concept inherent to human beings for it refers to the capacity of humans to think, choose and decide for themselves, implying self-awareness, self-consciousness, and self-authorship, it has been used to refer to the increasing degrees of automation and independence of machines from human control in terms of their operation and decision procedures.[80] When applied to the current state of AI, *autonomy* refers to the functional capacity of an artificial agent to operate independently from the direct control of human operators[81] and make decisions based on an evaluation of their options,[82] which resumes to their capacity of perceiving their environment and adapting their behaviour accordingly.[83] To some scholars as most of the autonomous systems that surround us still require human intervention under a range of conditions (e.g., an intelligent vacuum cleaner that falls down the stairs or is trapped

[76] FLOWERS, Johnathan – "*Strong and Weak AI: Deweyan Considerations" in* AAAI Spring Symposium: Towards Conscious AI Systems, 2019, p. 2.
[77] HAENLEIN, Michael; KAPLAN Andreas, – "*Siri, Siri, in my hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence* in Business Horizons", 2019, p. 16.
[78] Despite the several available definitions of AI, the varying level of autonomy is a common feature of some of them: e.g., COM (2018) Artificial Intelligence for Europe, p. 1. *See* also: OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, p. 7: " (…) AI systems are designed to operate with varying levels of autonomy."
[79] See PRIEST, Colin – "*Humans and AI: Should we describe AI as autonomous?",* 2021. Available at Blog / AI & ML Expertise: https://www.datarobot.com/blog/humans-and-ai-should-we-describe-ai-as-autonomous/
[80] European Group on Ethics in Science and New Technologies – "*Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems",* 2018, p. 9.
[81] TOTSCHNIG, Wolfhart – "*Fully autonomous AI"* in Science and Engineering Ethics, n.º 26(5), 2019: "In the field of artificial intelligence and robotics, the term "autonomy" is generally used to mean the capacity of an artificial agent to operate independently of human guidance." *See also*: BEER, Jenay M; FISK, Arthur. D; ROGERS, Wendy. A – "*Toward a framework for levels of robot autonomy in human-robot interaction"* in Journal of Human-Robot Interaction, 2014, p. 76.
[82] GLESS, Sabine, SILVERMAN, Emily, WEIGEND, Thomas – *"If Robots Cause Harm, Who is To Blame? Self-Driving Cars and Criminal Liability",* 2016, p . 442. And European Group on Ethics in Science and New Technologies - *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems,* 2018, p. 9.
[83] European Group on Ethics in Science *supra note* 85. *And* MSI-AUT(2018)05. Committee of Experts on Human Rights Dimensions of Automated Data Processing and different forms of Artificial Intelligence - *A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework.* 2018, p.13.

in a corner will need human intervention to resume its course; even automated vehicles which are the vanguard technology regarding automation still require human supervision) they should be defined as semi-autonomous systems (SAS).[84]

That said, we are not referring to full (self-deterministic) autonomy, which would only be feasible with strong AI, we are instead referring to autonomy as a range property that may be more or less present in degrees,[85] depending on the extent to which human intervention and oversight are necessary for a system to operate.[86] Several classification systems, taxonomies, and models have been proposed to evaluate the different levels of system autonomy. The most recent model for types and levels of automation provides a framework where the different functions of a system can be automated to differing degrees in a continuum of low to high (e.g., fully manual to fully automated), and different stages of automation represent input and output functions: information acquisition; information analysis; decisions and action selection; action implementation.[87]

To achieve autonomy, machines rely upon another distinctive trait of AI: machine learning algorithms. Machine Learning (ML) that "allows systems to learn directly from examples, data and experience."[88] ML focuses on the use of data and algorithms to imitate the learning process of humans, gradually improving its accuracy. These techniques have been thriving due to the increasing amounts of available data that have been generated during the last years as they shortly resume the capacity of an AI system to directly learn and improve itself from collected data.[89]

---

[84] ZILBERSTEIN, Shlomo – "*Building Strong Semi-Autonomous Systems" in* Twenty-Ninth AAAI Conference on Artificial Intelligence, vol. 29, n.º 1, 2015, pp. 4088, 4089.

[85] PARASURAMAN, Raja; SHERIDAN, Thomas. B; WICKENS, Christopher D. – "*A Model for Types and Levels of Human Interaction with Automation*. in IEEE Transactions on Systems, Man, and Cybernetics-Part A: Sistems and Humans", vol. 30, n.º 3, 2000, p. 287.

[86] MSI-AUT(2018)05. *supra note* 86 at 14.

[87] This model is proposed by Raja Parasuraman, Thomas Sheridan, and Christopher D. Wickens *supra note* 88 at 4, 5. *See also:* BEER, Jenay. M; FISK Arthur. D; ROGERS Wendy. A – *Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction.* In J Hum Robot Interact, 2014, pp. 4-5.

[88] The Royal Society Report - *"Machine Learning: The Power and Promise of Computers that Learn by Example"*, 2017, p. 16. Available at: https://royalsociety.org/-/media/policy/projects/machine-learning/publications/machine-learning-report.pdf

[89] IBM Cloud Education – *What is Machine Learning.* 2020. Available at: https://www.ibm.com/cloud/learn/machine-learning and OLIVEIRA, Arlindo – "*Inteligência Artificial"*, 2019, pp. 12-13. See also GIUFFRIDA, Iria – *Liability for AI Decision-Making: Some Legal and Ethical Considerations*. 2019, p. 441; BORGESIUS, Frederik Zuiderveen – *"Discrimination, Artificial Intelligence, and Algorithmic Decision Making",* 2018, p. 13

Designed to automate the learning process of the algorithms behind the machine's decision, with these techniques the knowledge in the system does not have to be provided by experts, as the system learns how to achieve the desired goals by itself. [90] ML allows for algorithms to constantly adapt and improve themselves based on the collected data while in use,[91] in this way, algorithms are able to infer certain patterns based on a set of data with more precision and to determine autonomously the best actions to achieve a specific goal.[92] These techniques are already found in several applications of our daily lives; speech recognition, facial recognition, auto-correct on smartphones, email spam detection, and automated driving are examples of deployed machine learning.

The basic function of ML techniques involves the use of statistical learning and optimization methods that allow computers to analyze datasets and identify patterns while leveraging on *data mining*[93] to identify historic trends and inform future models.[94] The process may be briefly resumed in three parts: a *decision process*, as generally these algorithms are used to make predictions and classifications. Based on input data (labeled or unlabeled) the algorithm must produce an estimate about a pattern in that data; *an error function* that serves to evaluate the prediction of the model and make comparisons to assess the model accuracy (this is a way of measuring how good was the algorithm guess by comparing it with the known examples); and *an updating or model optimization* process, through which the algorithm will repeat and optimize the process by updating weights autonomously until a threshold of accuracy is achieved.[95]

---

[90] SAMOILI, Sofia, et al. – *AI Watch Defining Artificial Intelligence: Towards an operational definition and taxonomy of artificial intelligence*. JRC Technical Reports. 2020, p. 12. *See also* BORGESIUS, Frederik Zuiderveen – Discrimination, Artificial Intelligence, and Algorithmic Decision Making. 2018, p. 13

[91] LIPTON, Zachary C. – *From AI to ML to AI: On Swirling Nomenclature & Slurried Thought*. 2018. *And* COM (2020) 65 final. Brussels, 19.02.2020 – *White Paper on Artificial Intelligence – A European approach to excellence and trust*, p. 16

[92] BOUCHER, Philip – "Artificial Intelligence: How does it work, why does it matter, and what can we do about it", 2020, Study for the Panel for the Future of Science and Technology, p.3. *And* BOSTROM, Nick – "*Superintelligence: Paths, Dangers, Strategies"*,2014, p. 23

[93] NILSSON, Nils J. – *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. 2010, p.500: "Data mining is the process of extracting useful information from large databases."

[94] Berkeley School of Information. *What is Machine Learning (ML)?.* 2020 available at: https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/

[95] IBM Cloud Learn Hub. Artificial Intelligence. *What is Machine Learning? .* 2020 available at https://www.ibm.com/cloud/learn/machine-learning#toc-challenges-L8chLUzD and Berkeley School of Information. *What is Machine Learning (ML)?.* 2020 available at: https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/

There are three main types of learning techniques:[96]

First, supervised learning methods are the most reliable and safe ones, as they refer to the use of pre-labeled and pre-classified datasets to train algorithms to classify data or predict outcomes accurately.[97] The algorithm learns to relate a given set of inputs to a given set of outputs (e.g. from a given data set of pictures containing cats and dogs, the system must be able to train itself to identify correctly each species)[98]

Second, unsupervised learning, where the algorithm is used to analyze and cluster unlabeled datasets. The algorithm explores input data without receiving an explicit output variable, the algorithm will discover the underlying hidden patterns and data groupings without human intervention (e.g., determine customer demographic data to identify clusters of consumer patterns; recommender systems).[99]

Third, reinforcement learning is a behaviour model where the algorithm develops by making sequences of decisions under different conditions through trial and error and maximizing the output when receiving validation.[100] Reinforcement learning is less secure as it is not possible to assess the accuracy or correctness of the resulting output. They require greater trust and confidence in the algorithm.

Artificial Neural Networks (ANNs) and Deep Learning (DL) are some of the adopted algorithms in the use of the aforementioned ML techniques. ANNs are inspired by the electro-chemical neural networks of the human brain, it has as input the data coming from

[96] MANSON, Stephen; SENG, Daniel – *"Artificial Intelligence and Evidence"*. 2021, p. 245.
[97] HAENLEIN, Michael; KAPLAN Andreas, – *Siri, Siri, in my hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence* in Business Horizons 62. 2019, p. 19 and Berkeley School of Information. *What is Machine Learning (ML)?.* 2020 available at: https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/
[98] BOUCHER, Philip – Artificial Intelligence: How does it work, why does it matter, and what can we do about it. Study for the Panel for the Future of Science and Technology. 2020, p.4. and *What is Machine Learning (ML)?.* 2020 available at: https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/
[99] *Id.*
[100] HAENLEIN, Michael; KAPLAN Andreas *supra note* 105 at 19. *And* Berkeley School of Information. *What is Machine Learning (ML)?.* 2020 available at: https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/; Example available at: COM (2020) 65 final. Brussels, 19.02.2020 – *White Paper on Artificial Intelligence – A European approach to excellence and trust*, p. 16.

the sensors and as output the interpretation of the collected images, in between there are hidden layers that manipulate the signals and transform the networks so they can provide a response.[101] When a neural network is multilayered, Deep Learning is used to process such information. It is designed to process a wide range of data resources without previous human preprocessing, with more accuracy. The network will ingest vast amounts of data and process them through multiple layers that learn increasingly complex features of the data at each layer.[102]

Learning systems distinguish themselves by their capability to learn and "dynamically setting their intermediate sub-goals and adapt to local conditions according to the collected or inputted data without human intervention", therefore their actions are not fully deterministic or predictable due to the wide variety of contexts and environments in which they operate.[103] This dynamism and interactions may result in an unexpected or not clear decision, which raises the problem of inscrutability and opacity that will be discussed when exploring the admissibility of AI-generated evidence in the Portuguese criminal system (see section 4.).

### 2.2 The attempt of defining AI at EU level

In order to achieve one of its core goals of an internal market that amongst sustainability, competitiveness, full employment, and social progress, the EU must also promote scientific and technological advances,[104] and be aware of the new trends that might impact citizens' lives and economy.[105] In 2018, the European Commission (EC) in the communication "AI for Europe" recognized AI as "one of the most strategic technologies of the 21st century"[106] launching the guidelines to construct a European initiative on AI,

[101] The European Commission's High-Level Expert Group on Artificial Intelligence – *A Definition of AI: Main Capabilities and Scientific Disciplines*. 2018, p.4. and BOUCHER, Philip – Artificial Intelligence: How does it work, why does it matter, and what can we do about it. Study for the Panel for the Future of Science and Technology. 2020, p.4.

[102] An Executive's guide to AI: *Deep Learning* available at: https://www.mckinsey.com/business-functions/quantumblack/our-insights/an-executives-guide-to-ai

[103] European Group on Ethics in Science and New Technologies - *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems.* 2018, p. 9.

[104] Article 3º/3 Treaty on European Union (TEU); Article 26º and 114º Treaty on the Functioning of the European Union.

[105] See COM (2017) 228 final. Brussels, 10.05.2017 – *A Connected Digital Single Market for All* , p.2.

[106] COM (2018) 237 final. Brussels, 25.04.2018 – *Artificial Intelligence for Europe*, p. 1.

which among boosting EU's technological and industrial capacity, increase AI uptake across the economy and preparing for socio-economic changes brought by AI, envisaged as well to ensure an appropriate ethical and a legal framework based on EU's values and fundamental rights.[107]

All these considerations were reinforced with the publication in 2020 of the White Paper on Artificial Intelligence, where the European Commission recognized the growing importance of AI as part of the "evolving family of digital technologies" in high-impact sectors and its socio-economic benefits, also emphasizing AI's potential risks for fundamental rights and the need of a regulatory framework to achieve an ecosystem of excellence and trustworthy AI (see section 5).[108] It was in the aftermath of the White Paper on AI that stakeholders, experts, and scholars waited with great expectations for a Regulation on AI, as it could dictate the success or failure of Europe's Digital Single Market project and EU's leadership in a digital economy.[109]

The EC published in 2021 the Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts (Artificial Intelligence Act).[110] The main objective of the Proposal is to present a balanced and proportionate horizontal regulatory approach, that while promoting the uptake of AI and encouraging the development of AI technologies simultaneously addresses its risks.[111]

To achieve that goal, the Proposal provides a technology-neutral and future-proof AI definition while adopting a risk-based approach, establishing prohibitions and different requirements and obligations according to the evaluated risk of each AI application.[112]

Despite having welcomed the Proposal initiative to regulate AI, experts are criticizing and recommending a revision of the proposed definition for AI:[113]

---

[107] DALLI, Hubert – *Briefing – Artificial Intelligence Act.* PE 964.212. 2021, p.1.
[108] COM (2020) 65 final. Brussels, 19.02.2020 – *White Paper on Artificial Intelligence – A European approach to excellence and trust*, p. 2, 3, 9.
[109]RAPOSO, Vera Lúcia – *Draft Regulation on Artificial Intelligence: The devil is in the details*. Privacy and Data Protection Magazin2. Nº3, 2021, p.11
[110] COM (2021) 206 final. Brussels, 21.04.2021 - *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts (Artificial Intelligence Act)*
[111] *Id* at 3.
[112] *Id.*
[113] See SMUHA, Nathalie, et al – *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal For an Artificial Intelligence Act*. LEADS Lab University of Birmingham.

> "Artificial intelligence system (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."[114]

The proposed regulation aims to be as future-proof as possible being able to accommodate future technological developments, to achieve such a goal, the definition is complemented by a list of techniques that usually constitute AI, such as machine learning, logic/ knowledge-based, and statistical approaches. Those techniques are listed in a separate annex which should be amended and updated as new technological advances emerge, through the adoption of delegated acts as offered in article 4.[115]

The provided definition is considered overly broad, potentially leading to a lack of clarity as well as to overregulation regarding high-risk AI systems making software that usually is not considered AI falling into the scope of this Regulation. Furthermore, the goal of achieving a future-proof definition is considered impossible due to the *odd paradox* that is natural from the rapid AI development.[116]

According to Robotics and AI Law Society (RAILS) experts, the definition provided in conjunction with Annex I techniques is prone to cover almost every computer program, which might generate uncertainty between AI developers and users "that associate AI primarily with machine learning, and not with simple automation processes in which pre-programmed rules are executed according to logic-based reasoning.[117] Besides, the main AI characteristics that endanger fundamental rights, such as opacity, complexity, and autonomy derive especially from machine learning techniques and not so much from

---

2021, p.2, 3, 14. *And* EBERS, Martin, et al - *The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)*. Multidisciplinary Scientific Journal. 2021, p. 499, 590.

[114] *Supra note 66.* at 39.

[115] COM (2021) 206 final. Brussels, 21.04.2021 - *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts (Artificial Intelligence Act),* p. 12, 43.

[116] RAPOSO, Vera Lúcia – *Draft Regulation on Artificial Intelligence: The devil is in the details*. Privacy and Data Protection Magazin2. Nº3, 2021, p.13

[117] EBERS, Martin, et al – "*The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)"* in Multidisciplinary Scientific Journal, 2021, p. 590.

simple logic-based algorithms.[118] Sharing the same opinion, Legal, Ethical & Accountable Digital Society (LEADS) members similarly understand that the description of software provided is incredibly broad as it encompasses virtually all algorithms, and the techniques listed in the annex include all computational techniques, being hard to determine if the regulation applies to computer scientists working with any technique draw on logic or statistical insight that is not conventionally seen as AI.[119]

In fact, as previously explained (see Section 2.), some of the techniques listed in Annex I – symbolic reasoning and expert systems – are not considered part of the second (current) wave of AI technology which is characterized by increased levels of automation and learning techniques.[120] These systems assume that human intelligence can be formalized and encoded in an "if-then-else rule" format, following the example of a symbolic AI system for medical purposes: "*If* the patient has a fever *then* prescribe drug X. *Else* send the patient home" or " *If* the patient has a fever and is allergic to drug X, *then* prescribe drug *Z*".[121] These systems are limited to constraint environments and the evolving variables are unambiguous and quantifiable. They tend to perform poorly on tasks that depend on complex forms of reasoning that cannot be translated into simple rules. Hence an expert system by itself is not considered an AI system.[122]

Considering the above critiques, two recommendations are made by the experts, the first passing by broadening the scope of the Proposal and changing its name to "Algorithm

---

[118] *Id.*

[119] SMUHA, Nathalie, et al – "*How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal For an Artificial Intelligence Act".* LEADS Lab University of Birmingham, 2021, p. 14, 15. About this matter see also RAPOSO, Vera Lúcia – "*Draft Regulation on Artificial Intelligence: The devil is in the details"* in Privacy and Data Protection Magazine, n. º3, 2021, p.12, 13.

[120] HAENLEIN, Michael; KAPLAN Andreas, – "*Siri, Siri, in my hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence* in Business Horizons", 2019, p. 18.

[121] This example may be found in BOUCHER, Philip – "*Artificial Intelligence: How does it work, why does it matter, and what can we do about it"*. Study for the Panel for the Future of Science and Technology, 2016, p.2.

[122] HAENLEIN, Michael; KAPLAN, Andreas –*" A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence"*, 2019, p. 4: "Expert Systems perform poorly in areas that do not lend themselves to such formalization.(…) an Expert System cannot be easily trained to recognize faces or even to distinguish between a picture showing a muffin and one showing a Chihuahua. For such tasks it is necessary that a system is able to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation—characteristics that define AI. Since Expert Systems do not possess these characteristics, they are technically speaking not true AI".

Act or Software Act", the second alternative recommendation is to limit the scope of AI intrinsic techniques to only include systems that rely on machine learning methods.[123]

Taking into consideration this criticism and the state of the art of AI technology the definition that will be used for the purposes of the following analysis is the one provided by the HLEG.[124] As explained before, this definition consists of a technical definition while referring to the intrinsic aspects and relevant techniques behind AI, focusing on the main techniques behind today's AI, namely machine learning.

In sum, AI is a collection of technologies that combine computing power, data, and algorithms. The present criteria to define AI technology consists of considering its intrinsic aspects and elements, focusing on a working definition instead of pursuing the concept of human intelligence. AI consists of software systems designed by humans, which act in the digital or in the physical world (through hardware) with learning rationality, to achieve a specific goal. In order to achieve rationality, these systems must be able to perceive their environment, collect and interpret data, adapt and reason on the knowledge obtained from the collected data deciding the most suitable action to take. Such capabilities are enabled by reasoning/decision-making techniques and machine learning. The current stage of scientific evolution is "limited" to narrow AI, non-self-conscient systems designed to perform specific tasks autonomously, without or with little human intervention.

## 3. AI at the service of criminal justice

Law is not an exception as regards the impact of AI, as, by its nature, it reflects the socio-cultural and economic context of the society in which it is inserted. "Law changes as its constraints change"[125] (*Ubi societas ibi ius*).

Criminal law as "the ultimate reaction of a jurisdiction to the aggression upon the core values of the society, (…) is strictly embedded in the social culture", it adapts to new

[123] SMUHA, Nathalie, et al – "*How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal For an Artificial Intelligence Act".* LEADS Lab University of Birmingham. 2021, p.15

[124] As detailed previously on section2.2; The European Commission's High-Level Expert Group on Artificial Intelligence – *A Definition of AI: Main Capabilities and Scientific Disciplines*. 2018, p.7.

[125] KARNOW, Curtis E.A in Research Handbook on the Law of Artificial Intelligence, 2018, p. xviii

realities and social needs as it "crystallizes accomplished processes into sets of commands, reflecting an accepted framework of social values". [126]

The digitalization and automation of bureaucratic tasks in the realm of legal professions benefited justice as strenuous handwritten and manual steps were replaced by digital platforms and software (e.g., since 2009 the Portuguese Justice System has implemented CITIUS[127], a digital platform that englobes informatic applications, the main objective of which is to dematerialize judicial procedures and promote procedural management and efficiency, by allowing Judicial Magistrates and Attorneys-at-Law to have digital access and digitally submit pleadings and supporting documents).

Nevertheless, the way AI serves criminal justice goes far beyond the benefits associated with the digitalization of the legal profession. If a few years ago criminal law had to adapt to the digital revolution by acknowledging a new type of crime – cybercrime - committed in a new (digital) environment[128], the impact of AI is much wider as it may affect the way justice is delivered.[129] The increasing ability of machines to perceive their surroundings, make autonomous decisions and predictions, as well as interfere with the physical environment through action, have turned intelligent systems into appealing law enforcement instruments and means of generating evidence.[130]

The access to unprecedented quantities of data coupled with the capability of processing the acquired data when confronted with a specific context allows AI systems to make a series of predictions and assertions that, when applied to criminal justice settings,

[126] *Id* at 1520.
[127] CITIUS means "fast" in latin, it is an informatic platform created to dematerialize and promote justice efficient, see website: https://www.citius.mj.pt/portal/faq.aspx
[128] Budapest Convention on Cybercrime and in the Portuguese context: *"Lei do Cibercrime"*, Law n.º 109/2009.
[129] Quattrocolo Supra *note* 132 at 1522-1523.
[130] RODRIGUES, Anabela Miranda – "*Inteligência Artificial no Direito Penal – A Justiça Preditiva entre a Americanização e Europeização* in A Inteligência Artificial no Direito Penal", 2020, p. 12. See also PAGALLO, Ugo; QUATTROCOLLO, Serena – "*The Impact of AI on Criminal Law and its Twofold Procedures* in Research Handbook on the Law of Artificial Intelligence", 2018, p. 386, 397: "AI technology can indeed be used for law enforcement purposes, or for committing (new kinds of) crimes"; "Our thesis is that AI systems, which will increasingly be used to generate evidence within criminal proceedings, entail a new set of issues that concern matters of transparency group profiling, loss of confidentiality, and more".

translate into what is currently defined as algorithmic justice or automated criminal justice.[131] In this context, AI systems started to pervade the realm of criminal justice through sophisticated risk-assessment tools or predictive systems designed to assist judges in the decision making process.[132]

The use of predictive justice systems, also known as "actuarial methods", is a common phenomenon in the USA,[133] they are designed to predict future criminal behaviour through the analysis of potential risk factors calculated from large datasets associated with specific traits of the defendant.[134] These systems may assist decision-making in different stages of criminal proceedings: since pre-trial detention and bail, to recidivism and escape risk calculation during assessment of culpability, and at the sentencing phase regarding parole and early release decisions.[135]

According to Harcourt[136], the majority of US-American scholars recognizes that predictive systems enhance justice efficiency and are beneficial to the society. These types of methods also find their way in American Law.[137] Besides that, some relevant associations such as the American Bar Association, the National Association of Counties, the Conference of Chief Justices and the National Center for States Court positioned themselves in favour of the use of risk assessment tools in pretrial and sentencing

[131] ZAVRŠNIK, Aleš – *"Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings",* 2019, p. 2-4. And *SACHOULIDOU, Athina – "OK Google: is (she) guilty?"* in Journal of Contemporary European Studies, 2021, p. 1.

[132] *Id* at 3. *See also*: PAGALLO, Ugo; BARFIELD, Woodrow – "*Advanced Introduction to Law and AI",*2020 p. 10,11.

[133] GIALUZ, Mitja – "Quando la Giustizia Penale Incontra L'Intelligenza Artificiale: Luci e Ombre Dei Risk Assessment Tools Tra Stati Uniti Ed Europa"in Diritto Penale Contemporaneo, 2019, p. 4.

[134] HARCOURT, Bernard. E - *"Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age". 2015,* p. 4*: "*Actuarial methods consist of the use of statistical rather than clinical methods on large datasets of criminal offending rates to determine different levels of offending associated with one or more group traits in order to predict past, present or future criminal behavior and administer a criminal justice outcome."

[135] LIGETI, Katalin – *"Artificial Intelligence and Criminal Justice"* in AIDP-IAPL International Congress of Penal Law, 2019, page 7; *SACHOULIDOU, Athina – "OK Google: is (she) guilty?"* in Journal of Contemporary European Studies, 2021, p. 3.

[136] HARCOURT, Bernard E – Against Prediction: Sentencing, Policing and Punishing in an Actuarial Age , May 2005, pp. 14–15, available at https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1021&context=public_law_and_legal_theory.

[137] The Unites States Model Penal Code in 2017 exhorted the use of actuarial instruments. GIALUZ, Mitja – "Quando la Giustizia Penale Incontra L'Intelligenza Artificiale: Luci e Ombre Dei Risk Assessment Tools Tra Stati Uniti Ed Europa"in Diritto Penale Contemporaneo, 2019, p. 4.

phases.[138] Nevertheless, the current implementation of predictive justice systems to criminal justice in Europe is very rare, with some exception as for HART (Harm Assessment Risk Tool) implemented in UK, which objective is to determine the recidivism risk of detained persons during a period of two years of detained people.[139] The use of AI in continental European justice systems remains primarily under the private-sector commercial initiative aimed at legal departments, lawyers, insurance companies and individuals (eg., online alternative dispute resolution, chatbots to inform litigants and provide support in legal proceedings, and advanced case-law engines).[140]

The use of predictive justice systems at the service of criminal justice has been associated with high expectations of increased criminal justice efficiency and accountability, leading to more accurate knowledge and rationalization of human performance.[141] Nevertheless, this comprises a high risk to fundamental rights mainly on the perspective of the defence rights as raised from *Loomis v. Wisconsin* case, where the sentencing was based in the result provided by a risk assessment tool: COMPAS (Correctional Offender Management Profiling for Alternative Sanction). This tool builds a profile for the individuals based on their response to a set of elaborated questions, and criminal record analysis. Such information will be statistical weighted in accordance to grouped datasets.[142]

In 2013 Eric Loomis was charged for five crimes that were related to drive-by shooting, driving without license, escape attempt, illegal gun possession, and dangerous driving. He was convicted to maximum penalty for the practice of such crimes and didn't qualify for probation. During the trial the court referred the result provided COMPAS – *high risk of recidivism* - as a foundation for the decision. Loomis appealed the decision referring to the lack of transparency regarding the criteria behind the result, also referring the

[138] *Id*.
[139] RODRIGUES, Anabela Miranda – *"Inteligência Artificial no Direito Penal – A Justiça Preditiva entre a Americanização e a Europeização"* in A Inteligência Artificial no Direito Penal, 2020, p. 18, 19.
[140] European Commission for the Efficiency of Justice (CEPEJ) – *"European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment"*, 2018, p. 16.
[141] *SACHOULIDOU, Athina – "OK Google: is (she) guilty?"* in Journal of Contemporary European Studies, 2021, p. 3.
[142] *See* HARCOURT, Bernard. E - *"Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age",* 2015, p. 4. And CARIA, Rui – *"O Caso State v. Loomis – A Pessoa e Máquina na Decisão Judicial"* in A Inteligência Artificial no Direito Penal, 2020, p. 247, 248.

violation of his right to a fair trial.[143] As COMPAS was designed to analyze datasets grouped by gender and race also generated concern regarding the use of discriminatory algorithms to support criminal decisions.[144]

The use of these predictive systems challenges are similar to the ones that will be analyzed in the following chapters regarding AI-generated evidence, such as the right to a fair trial, the right to an effective defence and respect for the presumption of innocence, therefore this discussion will be addressed to the next sections. This analysis will focus from now on in AI generated machine evidence.

### 3.1 AI Generated Evidence

From a pretrial stage where inquiry measures are taken to determine if a real crime was committed and to identify the crime perpetrators, to the ruling regarding coercive procedural measures, such as pre-trial detention, until the end of the trial hearing and sentencing, evidence takes an essential role in the bearing of a proper decision, see article (art.) 124º Portuguese Code of Criminal Procedure (PCCP).[145] In accordance to Germano Marques da Silva, the essential function of a criminal procedure is to decide whether a crime was committed, determine its perpetrators and determine criminal liability.[146] To achieve such a goal evidence plays a crucial role, according to Bentham the criminal procedure is nothing more than the art of administrating evidence.[147]

According to the Portuguese law, the court must order the production of all essential means of evidence to achieve the truth and a good ruling of the case, art. 340º, n.º 1

[143] CARIA, Rui – *"O Caso State v. Loomis – A Pessoa e Máquina na Decisão Judicial"* in A Inteligência Artificial no Direito Penal, 2020, p. 248 – 254. See also: Harvard Law Review – *"State v. Loomis Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing.",* 2017. Available at: https://harvardlawreview.org/2017/03/state-v-loomis/

[144] ZAVRŠNIK, Aleš – *"Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings".* 2019, p. 3.

[145] ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem."* 2018, p.330.

[146] SILVA, Germano Marques – *"Curso de Processo Penal",* Tomo I, 2010, p. 35. And ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem",* 2018, p.329 - 330.

[147] Bentham APUD FERREIRA, Manuel Marques - *"Meios de Prova"* in Jornadas de Direito Processual Penal, Lisboa, 1988, pág. 221 a 260.

PCCP.[148] *Princípio da Investigação ou da Verdade Material*[149]: This is one of the leading principles of the Portuguese criminal procedure. According to this principle, the court has the duty to perform all the necessary diligence to assess the truth behind the facts, by demanding the presentation of any proof that appears to be of importance to reach a decision even if that proof has never been presented at a preliminary stage of the procedure. Scholarship[150] and jurisprudence[151] have considered this an ethical and democratic principle, that demands the assessment of a material truth, as a requirement for a fair justice.

Nevertheless, the assessment of truth is limited by the defendant's rights, the right to a fair trial in particular, and the accusatory structure of the Portuguese criminal procedure.[152] The right to a fair trial is a fundamental right enshrined in art. 6º of the European Convention of Human Rights (ECHR), and the accusatory structure as enshrined in art. 32º, n.º 5 of the Constitution of the Portuguese Republic (CPR) aim for a constant balance and proportionality between the prosecution, truth assessment and the defendants' fundamental rights.[153] From this, results that the Portuguese criminal procedure is led by an adversarial nature (*estrutura acusatória*), mitigated by an investigation principle,[154] meaning "there is no valid Criminal Procedure without sustaining evidence, nor a legitimate Criminal Procedure without respect for the defendant's safeguards".[155] Hence, the admissibility of new means of evidence must be guided by the need of an effective truth assessment, limited by the fundamental safeguards imposed by the Accusatory nature of the procedure.

---

[148] Art. 340º/1º/2º CPP.

[149] Regarding the lack of a literal translation to English it was decided to refer to this principle in the original language. It consists of the relevance of Investigation to access the truth behind the facts.

[150] MESQUITA, Paulo Dá – *"A Prova do Crime e o que Se Disse Antes do Julgamento" – Estudo sobre a Prova no Processo Penal Português à luz do Sistema Norte-Americano"*, 2011, p. 263. Also FERREIRA, Manuel Cavaleiro de - Curso de Processo Penal, vol. I, 1955, p. 49, *and* FIGUEIREDO DIAS, Jorge – "*Direito Processual Penal",* vol. *I,* 1974, p.72.

[151] Refer to Constitutional Courts Rulling: AC. TC nº 137/2002 and AC.TC nº 394/1989.

[152] VALENTE, Manuel Monteiro Guedes – *Processo Penal.* Tomo I. 2020. pp-37-38.

[153] RAMALHO, David Silva – *"Métodos Ocultos de Investigação Criminal em Ambiente Digital"*, 2017, p. 182.

[154] In accordance with Jorge Figueiredo Dias: FIGUEIREDO DIAS, Jorge – *"Direito Processual Penal",* vol. I, 1974, p. 61.

[155] Translated from the original in Portuguese: "Não existe um processo penal válido sem prova que o sustente, nem um processo penal legítimo sem respeito pelas garantias de defesa." PINTO, Frederico de Lacerda da Costa; BELEZA, Teresa Pizarro - *"Prova Criminal e Direito de Defesa"* in Estudos sobre a Teoria da Prova e Garantias de Defesa em Processo Penal. 2013, p.5.

Humans are now surrounded by machines that can perceive their surrounding environment, actively surveil, collect data, act autonomously, and convey messages in response to human conduct[156], which may provide relevant information for truth assessment when confronted with the context of a criminal proceeding, creating what we refer to as *machine evidence*.[157]

The more AI systems pervade our lives, the higher is the possibility for them to "intrude" criminal proceedings and the need to examine whether and how to accommodate them in criminal legal frameworks.[158] Machine evidence as a result of AI-generated data and inferences has the potential to provide new sources of information, representing a chance for more accurate fact-finding and truth assessment in criminal trials.[159]

It is expected that with the emergence of AI, courts will face the question of whether it will generate reliable, accurate, and objective evidence.[160] In fact, this process is not something new, as since the turn of a new century the *crescendo* of digital technology and scientific knowledge revolutionized forensic and non-forensic science bringing new tools for evidence collection and production in criminal proceedings. The cycle of evaluating new types of evidence, testing their credibility, and finding a balance between criminal prosecution, truth assessment, and the safeguard of fundamental rights in criminal proceedings is part of what is defined as the *evidentiary life cycle*[161] of new types of evidence.[162]

[156] The leading example will be advanced safety systems embedded in automated vehicles, that will be discussed in the following chapters.
[157] GLESS, Sabine – *AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials* in George Town Journal of International Law, vol.51, Nº2, 2020, p. 195.
[158] *Id* at 207: "As AI becomes more ubiquitous, and if such technology is deemed to be an accurate assessment of human conduct, more people may be willing to accept it as a reliable and trustworthy source of information. Despite this possibility, it remains unclear if and how such information would be admitted into a court of law."
[159] *Id.*
[160] GLESS, *supra note* 171 at 210.
[161] There might be a predictable life cycle for most new types of evidence: starting from *to new to be reliable,* it becomes *to new but subject to testing*, evolves then *to generally reliable but occasionally improperly applied* and the final stage *being blindly trusted*. See GLESS, supra note 157 at 215
[162] MURPHY, Erin – *The New Forensics: Criminal Justice, False Certainty, and The Second Generation of Scientific Evidence*. 2006, p.4: ""Thus, in this age of powerful and pervasive new forensic technologies, the criminal justice system must reckon anew with how it accommodates scientific evidence."

The fact is "our current models of criminal justice, even operating at their optimal level, cannot adequately safeguard the widespread use of highly probative and sophisticated evidence".[163] Each evidence innovation is feasible of raising specific concerns about the need for adequate integration in criminal proceedings, and as more sophistication and innovation enter the courts, the more new challenges are raised.

This observation leaded Erin Murphy to present a taxonomy of evidence, distinguishing between the first and the second-generations of forensic evidence:[164]

The first-generation of forensic evidence is limited to specific categories of offenses (e.g., ballistics only applies to cases involving a gun; fingerprint collection), and they require the identification of an individual person or object for comparison (e.g., resuming the previous example of ballistic and fingerprint collection, even though a gun is found this technique is only feasible if a bullet is recovered or if the fingerprint is in a good state to be used). Also, the recovery rate of these residues is very low as they are difficult to find and preserve due to their delicacy (e.g., hair or tissue fiber strands). Therefore, "first-generation forensic sciences lack a robust investigative capacity to identify a suspect in the first instance, and instead operate mainly to confirm the defendant's connection to a crime"[165], this means they often rely on the support of other forms of evidence, such as eyewitness testimony. Also, first-generation evidence does not depend on technically sophisticated concepts, nor depend on complex machinery being intuitively comprehensible by laypeople. The last distinctive factor pointed out by Murphy is that first-generation scientific evidence does not implicate the same questions of personal privacy, protection, and proprietary information as second-generation scientific evidence.[166]

By contrast, the second-generation evidence does not have an offense-specific character, as they apply to a broad range of charges and case types. Their recovery rate is higher than first-generation evidence and can make them render irrelevant (e.g., where there is a fingerprint or hair evidence, there is often sufficient genetic material to conduct DNA

[163] *Id.*
[164] *Id* at 5.
[165] *Id*.
[166] *Id* at 6.

typing which is more reliable). Second-generation evidence offers stronger scientific certainty, as derive from technically sophisticated methodologies that require expensive equipment and particular expertise (e.g., DNA typing), they also have the potential to generate conclusive proof in the absence of other evidence and their specificities raise concern regarding privacy and exposure of proprietary information.[167]

The second-generation evidence also englobes non-forensic technology, such as digital evidence (e.g., GPS/location tracking; data mining).[168] There are different definitions for digital evidence amongst jurisprudence and scholarship, which is generally (erroneously) referred to as electronic evidence. This confusion might result from a blind adherence to the lettering of the Budapest Convention on Cybercrime, that defines electronic evidence as "evidence that can be collected in electronic form of a criminal offense."[169]

According to Weir and Manson, electronic evidence is a generative term englobing all data that "is manipulated, stored or communicated through any man-made device, computer or computer system, or transmitted over a communication system, that has the potential to make the factual account of either party more probable or less probable than it would be without the evidence".[170] Electronic evidence englobes all data independently from being produced or stored in an analogue device, or in digital form, hence digital evidence must be seen as a form of electronic evidence.[171] Following Kerr and Manson, digital evidence is a specific term and consists of the evidence generated from data in digital (binary) form.[172]

This taxonomy also considers its distinctive elements, such as *immateriality*, as it is composed of a sequence of *bits* existing independently from the material support in which it is stored, and *volatility* due to the fact it may be destroyed, erased, or damaged, hence the gathering of this type of evidence must operate through appropriate tools and

[167] *Id* at 7.
[168] GLESS supra note 171 at 215.
[169] This reference to electronic evidence is found in the preamble and several provisions of Budapest Convention on Cybercrime (2001), art. 14º, 23º, 25º and 35º.
[170] FIDALGO, Sónia – *A Utilização de Inteligência Artificial no Âmbito da Prova Digital – Direitos Fundamentais (ainda mais) em perigo* in A Inteligência Artificial no Direito Penal. 2020, p. 133. see also Electronic Evidence Guide – A basic guide for the police officers, prosecutors and judges. 2020, p. 4.
[171] Amongst Portuguese scholarship: RAMALHO, David Silva – *"Métodos Ocultos de Investigação Criminal em Ambiente Digital"*, 2017, p. 100.
[172] KERR, Orin – "*Digital Evidence and The New Criminal Procedure"* in Colombia Law Review, vol. 105 2005, p. 279.

procedures.[173] Digital evidence may be generated and stored in the most common devices from our daily lives, such as computers, smartphones, digital cameras, compact disks, and memory cards.[174]

Modern technology evolves at increasing speed, and the digital evidence to which courts were still adapting, is becoming more complex as digital tools can now be AI-driven. AI embedded devices are constituted by specific singularities that make them different from first and second-generation evidence. According to Gless and Nutter, the underlying technologies behind DNA tests and traditional digital evidence are different from AI embedded technologies, so while the first and second generation of forensic evidence relied on human expertise (are person-based), the new type of evidence we are approaching is guided by source code and machine learning algorithms, capable of generating its own assertions. Therefore, due to its specific functioning[175] and the new raised evidentiary issues, machine evidence should be considered a third-generation type of evidence.[176]

Even though the presented taxonomy of generations of evidence focuses on the perspective of evidence gathering techniques. The same reasoning will be transposed to approach means of producing evidence *(meios de produção de prova).*

Whereas traditional forms of person-based evidence - testimony (art. 131º PCCP) and documentary evidence (article 164º PCCP) - have been considered mainstays in the criminal proceeding, new sources of revealing relevant information for fact-finding have been gradually fostered. The efficiency and pervasiveness of the technological advances have been creating the need to criminal proceedings to adapt to new realities.

---

[173] FIDALGO, Sónia – *"A Utilização de Inteligência Artificial no Âmbito da Prova Digital – Direitos Fundamentais (ainda mais) em perigo* in A Inteligência Artificial no Direito Penal", 2020, p. 134.

[174] About the distinction between electronic/digital evidence and digital evidence taxonomy amongst Portuguese scholarship please refer to: RAMALHO, David Silva – *"Métodos Ocultos de Investigação Criminal em Ambiente Digital"*, 2017, p. 102. See also Electronic Evidence Guide – "*A basic guide for the police officers, prosecutors and judges."* 2020, p. 4.

[175] For further detail on how AI systems function refer to section 2.

[176] GLESS supra note 171 at 211. and NUTTER, Patrick W. – "*Machine Learning Evidence: Admissibility and Weight."* in Journal of Constitutional Law, vol. 23, nº3, 2019, p. 922: "However, few have explored machine learning as a distinct species of machine evidence, distinct even from evidence produced using traditional computer programs, with its own vocabulary and unique set of issues."

Such as the increased use of electronic devices created the necessity of a specific category of evidence for mechanical reproductions (article 167º PCCP), it will be a matter of time "before litigators encounter creative opposing counsel who wishes to admit AI generated output into evidence",[177] and the criminal proceeding finds the necessity to adapt once more.[178]

It's simple to understand why AI-generated evidence may enter soon in the legal debate, it seems to benefit fact-finding and an efficient truth assessment during criminal trials.[179] In favour of machine generated evidence, Andrea Roth when comparing it to traditional human witnesses refers that "machines if well operating don't suffer for memory loss as humans do, they also don't exhibit character for dishonesty".[180] Although, along with the efficiency and accuracy of machine evidence, we enter a hazardous field for fundamental rights, mainly when it comes to defendants' rights. Several questions on "if" and "how" machine evidence will be admitted in courts of law are arising. Machine evidence is new and unique and does not fit the current human-centric evidentiary paradigm, therefore it will be necessary for legal systems to consider the possibility to adapt to the inherent changes.[181] How will AI-generated machine evidence admissibility get tested and integrated into the current legal framework, and which fundamental rights are at stake are the main questions to be addressed in turn by reference to the Portuguese legal system.

## 4. The admissibility of evidence generated by means of AI embodied monitoring systems in automated vehicles in the Portuguese criminal justice system

Automated or autonomous Vehicles (AVs) are an example of a fast-evolving AI technology that has been a bet of automotive engineering during the last years.[182] They

[177] NUTTER, Patrick W. – "*Machine Learning Evidence: Admissibility and Weight."* in Journal of Constitutional Law, vol.23, n.º3, 2019, p. 920
[178] *Id* at 924.
[179] GLESS supra note 171, at 207: "Machine Evidence like other forms of technology that came before, has the potential to provide new sources of information and provide more accurate fact-finding"
[180] ROTH, Andrea – "*Machine Testimony*", 2017, p.1979
[181] NUNN, Alexander – "*Machine Generated Evidence in The SciTech Lawyer: A publication of the American Bar Association" in* Science & Technology Law Section, vol 16, n. º 5, 2020, pp. 4, 5.
[182] ENZWEILER, Markus – "*The Mobile Revolution – Machine Intelligence for Autonomous Vehicles"*, 2015, p.1.

consist of motor vehicles embodied with automation systems that support and assist humans performing the driving task and are able to take control in specific situations.[183] In autonomous driving, algorithms are used in real time to collect data from the vehicle itself (e.g., speed) and from the whole environment through its sensors (e.g., road and weather conditions, signs, surrounding vehicles and pedestrians) to assess which direction, acceleration or speed the vehicle should take to reach a destination with safety.[184]

As previously mentioned, autonomous systems refer to varying levels of autonomy. The Society of Automotive Engineers (SAE) distinguishes 6 levels of on-road motor vehicle autonomy: from levels 0 to 2 human drivers are still in charge of the driving tasks and must constantly supervise and be engaged in the driving task. While levels 0-1 show none or little automation, except for single driver assistance tools at level 1 (e.g., lane centering or adaptive cruise control), at level 2 (Partial Driving Automation) the vehicles have automated combined capabilities such as acceleration and steering (e.g., lane centering and adaptive cruise control simultaneously), but the driver must remain engaged with the driving task. From levels 3 to 5 the driving task is taken by AI software systems. At level 3 of automation (Conditional Driving Automation) the driving task is handed to software and the driver is not fully required to monitor the surrounding environment, however, the human must be ready to take control at all moments and override. Level 4 (High Driving Automation) is the turning point to increased automation, the vehicle is capable to perform all the driving functions under certain conditions (e.g., when it is not raining or in limited areas – geofencing – with specific speeds limit), however, humans still have the option to take control of the vehicle. Level 5 (Full Automation) is the ultimate level of driving automation, at this stage, a human driver is not needed as the vehicle is capable to perform all driving functions under all conditions.[185]

---

[183] *See* GLESS supra note 202. Also according to SAE (Society of Automotive Engineers) automated driving refers to "motor vehicle driving automation systems that perform part or all of the dynamic driving task on a sustained basis"- SAE J3016 Recommended Practice: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, 2021.

[184] COM (2020) 65 final. Brussels, 19.02.2020 – *White Paper on Artificial Intelligence – A European approach to excellence and trust*, p. 16.

[185] See SAE International's blog- "*SAE Levels of Driving Automation Refined for Clarity and International Audience"*, 2021. And SAE J3016 Recommended Practice: *"Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles",* 2021. also POSZLER, Franziska; GEISSLINGER, Maximilian – "*AI and Autonoums Driving: Key Ethical Considerations"*, 2021, p. 2.

The current stage of automation on most on-road vehicles resides on level 2. Nevertheless, car manufacturers and software companies are currently allocating efforts on levels 3 and 4 research and development, with special prevalence to level 3 constituted by AI systems that focus on detection and behaviour prediction of other road users and perform subsequent decision making.[186] Some software companies already tested the roads with fully autonomous vehicles (e.g., Google's Waymo in 2015 and Tesla's automatic pilot), however there were registered some fatal accidents evolving self-driving vehicles. In 2018 an Uber's self-driving car was charged for fatally hitting a pedestrian. Most recently in 2021, a full automated Tesla crashed causing the death of its two passengers.[187] In conclusion, despite the current testing the current level of automated driving autonomy still lingers between level 2 and 3 demanding the intervention of an attentive human driver.

There is "a spectrum of technologies between driver-operated vehicles and autonomous road vehicles",[188] which intervention does not actively change or eliminate the role of the driver in part or all of the driving task. Hence, they are not considered part of SAE's automation taxonomy.[189] They might play an important role in preventing road accidents and generate relevant conveyances for criminal justice purposes. AI embodied monitoring systems in automated vehicles with a focus on detecting drivers' inattention and

[186] POSZLER; GEISSLINGER supra note 183 at 2.

[187] JANUÁRIO, Túlio Xavier – "*Veículos Autónomos e Imputação de Responsabilidades Criminais por Acidentes* in A Inteligência Artificial no Direito Penal", 2020, p. 97. See also JONES-CELLAN, Rory – *"Uber's self-driving operator charged over fatal crash",* 2020, BBC news available at: https://www.bbc.com/news/technology-54175359
ISIDORE, Chris – "*Police say no one was in driver's seat in fatal Tesla crash* in CNN Business", 2021. Available at : https://edition.cnn.com/2021/04/19/business/tesla-fatal-crash-no-one-in-drivers-seat/index.html

[188] Royal Academy of Engineering Report – "*Autonomous Systems: Social, Legal and Ethical Issues",* 2009, p. 5.

[189] SAE J3016 Recommended Practice: "*Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles",* 2021: "Due to the momentary nature of the actions of active safety systems, their intervention does not change or eliminate the role of the driver in performing part or all of the DDT, and thus are not considered to be driving automation, even though they perform automated functions. In addition, systems that inform, alert, or warn the driver about hazards in the driving environment are also outside the scope of this driving automation taxonomy, as they neither automate part or all of the DDT, nor change the driver's role in performance of the DDT."

drowsiness[190] will be the leading example for the rest of the analysis, as currently, they are the strongest example of how AI technology can monitor human behaviour.

As automated driving has been considered a key element of EU's strategy to implement road safety,[191] it is expected to see these systems as part of the common traffic in EU roads. In accordance, the European Parliament recognized the potential of automated vehicles to reduce road fatalities by publishing in 2019 a new Regulation on Type-Approval Requirements for Motor Vehicles. The latter designates a set of mandatory advanced safety systems that must start integrating motor vehicles circulating in the EU roads starting in July 2022.[192] Amongst them, this analysis focuses particularly on the use of *advanced driver drowsiness and attention warning systems* and *advanced driver distraction warning*.

According to the definitions provided in art. 3º of the Regulation (EU) 2019/2144, driver drowsiness and attention warning "means a system that assesses the driver's alertness through vehicle systems analysis and warns the driver if needed" and advanced driver distraction warning "means a system that helps the driver to continue to pay attention to the traffic situation and that warns the driver when he or she is distracted".[193] These systems may be used at lower levels of automation when the driver still maintains the control of the vehicle as an advanced safety tool, and at levels of increased autonomy (level 3 in especial), generating Take-Over Requests (TOR). Although automated vehicles can perform the driving task to a large extent, there are situations when human drivers must take action or override the automated system. The overriding may take place by the human driver's own initiative or be triggered by a warning alert generated by the

[190] See DONG, Yanchao, et al –*"Driver Inattention Monitoring System for Intelligent Vehicles: A Review in IEEE Transactions On Intelligent Transportation Systems",* vol. 12, n.º 2, 2011, p. 596, 597.
[191] COM (2018) 293 final. Brussels, 17.5.2018 - *EUROPE ON THE MOVE - Sustainable Mobility for Europe: safe, connected, and clean,* p. 5 : "Driverless vehicles and advanced connectivity systems should make vehicles safer and easier to share, and open up access to mobility services for more users. These technologies can also help to address many of the major challenges facing today's road transport system, such as road safety (…)" *see also* Regulation (EU) 2019/2144 of 27 November 2019, (23): "Automated vehicles have the potential to make a huge contribution to reducing road fatalities, given that more than 90 % of road accidents are estimated to result from some level of human error".
[192] *See* Regulation (EU) 2019/2144 of 27 November 2019 on Type-Approval Requirements for Motor Vehicles and their Trailers, and Systems, Components and Separate Technical Units Intended for such Vehicles, as Regards their Safety and the Protection of Vehicle Occupants and Vulnerable Road Users, article 6º and article 19º.
[193] *Id,* article 3º.

system generating a TOR (usually represented by an acoustic gong and on-screen warning in the vehicle pane, it might also send the alert to the driver's mobile phone and/or produce an additional brake jerk), in that case, the human must assume the driving task within a time framework.[194]

Advanced safety systems or advanced driving assistants consist of "software bots[195] designed to enhance driving safety by observing and assessing a human driver's behaviour" by issuing an alert when the driver appears drowsy or distracted. They differ from traditional digital tools and devices, as they are embedded with a degree of autonomy that allows them to convey their own messages after processing the collected data related to the driving environment and the driver's conduct.[196]

As any other AI system, advanced safety systems use ML techniques, such as ANNs[197] and other nonlinear modeling techniques – Fuzzy Inference Systems (FIS) and Support Vector Machine (SVM) - to analyze the extracted data and generate an output result.[198] Currently, there are five main types of measures for inattention or drowsiness detection in use: subjective report measures; driver biological measures; driver physical measures; driving performance measures and hybrid measures. The latter is considered the most reliable to use in a driving context for combining driver physical measures with road scene information measures.[199] There are some companies that have been developing and commercializing driver inattention and drowsiness systems based on hybrid measures: Toyota developed for their Lexus models a Driver Monitoring System (DMS), the system is integrated with a camera, which uses near-infrared technology, on top of the steering

[194] MELCHER, Vivien, et al. – "*Take-Over Requests for Automated Driving."* In Procedia Manufacturing, n.º 3, 2015, p. 2868-2886.

[195] Bot is here understood as an automated software interface that connects human users and other IT systems to services, such services might be internalized in the bot's code or accessed externally, according to the presented taxonomy in LEBEUF, Carlene R. - *"A Taxonomy of Software Bots: Towards a Deeper Understanding of Software Bot Characteristics"*. 2018, p. 17-20.

[196] GLESS, supra note 157 at 204, 205.

[197] Already mentioned in chapter 2.1;

[198] FIS distinguishes itself for its linguistic concept modeling ability. The fuzzy rule "is close to an expert natural language", it manages uncertain knowledge and infers high-level behaviours from the observed data. SVM is based on statistical learning techniques for pattern classification and "inference of nonlinear relationships between variables", it is a method used to detect and recognize faces, objects, text, speech, handwritten characters, and retrieves information and images. This learning method is more suitable to measure cognitive states of humans. See DONG, Yanchao, et al – *"Driver Inattention Monitoring System for Intelligent Vehicles: A Review."* in IEEE Transactions On Intelligent Transportation Systems, vol. 12, n.º 2, 2011, p. 601.

[199] *Id* at 601, 605.

column cover. The camera monitors the exact position and angle of the driver's head. In combination with the Advanced Precrash Safety system (APS), the vehicle's sensors detect obstacles ahead, if there is a near obstacle and the DMS determines that the driver's head is turned away from the road for a certain period, a precrash warning is triggered. Also, Toyota's Crown System can detect drowsiness by monitoring the driver's eyelids.

Another example is the Swedish company Saab which implemented the Driver Attention Warning System in its vehicles. This system is designed to detect both drowsiness and distraction, it uses two miniature infrared cameras (one camera is installed at the base of the driver's A-pillar and the other at the center of the main fascia), which must focus on the driver's eyes. The SmartEye software will analyze eyelid movements, gaze (glance behaviour and visual tracking), head orientation, and eye blinking frequency. If the software detects a pattern of long-duration eyelid closure, it identifies a potential context of drowsiness. This software is also monitoring if the driver's gaze moves away from the "primary attention zone" – the central part of the windshield – if within a time frame of seconds, the driver's eyes do not return to position, then the system triggers the alert for a distraction. The algorithm takes into consideration peripheral tasks and maneuvers that require the driver to change the focus from the central part of the windshield, in that case the timer's elapse time is longer. When the system detects drowsiness a three-level warning interface is triggered, it starts with a sound signal and text message, evolving to a spoken message, and ends in a stronger audio warning that must be reset by the driver. When distractions are detected, a vibration signal is sent to the driver's seat.[200]

In some vehicles (e.g., Honda and Mercedes-Benz) the drowsiness and distraction alert are illustrated by the sign of coffee cup icon and a 4-level bar graph on the car panel. When the bar drops to half, the sign is illuminated urging the driver to take a break. If the

[200] All these examples may be found at DONG, Yanchao, et al – *"Driver Inattention Monitoring System for Intelligent Vehicles: A Review."* in IEEE Transactions On Intelligent Transportation Systems, vol. 12, n.º 2, 2011, p. 599, 600. See also: HANSEN, John. H. L, et al. – "*Driver Modeling for Detection & Assessment of Distraction: Examples from the UTDrive testbed"*, 2017, p 2-11. and BERGASA, Luis M., et al. – *"Real-Time System for Monitoring Driver Vigilance"* in IEEE Transactions on Intelligent Transportation Systems, vol. 7, n. º 1.

driver continues driving until the bar drops to the lowest level an audible warning is triggered and the steering wheel may vibrate. [201]

Driving automation and the use of advanced safety systems exemplify "a gray area around the use of AI for evidentiary purposes".[202] As this technology becomes more ubiquitous and soon mandatory, it will be a matter of time until the machine's assessments about the driver's conduct will be tested as a reliable source of information and enter the courtroom.

According to the Portuguese law, driving without reuniting the necessary physical or psychic conditions to ensure safe driving, including under the influence of excessive fatigue, and thus endangering others' life, physical integrity or high value property incurs in the criminal offence of dangerous driving punishable with imprisonment up to three years according to art. 291º Portuguese Penal Code (PPC). What happens in the case of a car crash where the drowsiness warning signal is illuminated in the car's panel or when any other audible warning is active? It is still to be determined if and how such information could be admitted as evidence before Portuguese criminal courts.

### 4.1 The compliance of AI-generated evidence with criminal procedural rights

The first obstacle when assessing the admissibility of AI-generated machine evidence concerns to the principle of legality, and how to integrate this new type of evidence into the existing evidentiary framework in the Portuguese criminal justice system. According to the Portuguese Law, "all evidence is admissible when not prohibited by law,[203] art.125º PCCP. This means that, unlikely to some other jurisdictions (e.g., Italy), the Portuguese evidentiary framework is not restricted to the already predicate catalogue of evidence of articles 128º - 164º PCCP, materializing in this way the investigation principle and benefiting truth assessment.[204]

---

[201] Honda Website – *"Driver Attention Monitor"*, available at https://www.honda-mideast.com/en/technology/Driver-Attention-Monitor and Mercedes-Benz website - "The Mercedes-Benz ATTENTION ASSIST® System Can Detect and Alert You to Drowsy Driving". available at: https://www.mercedesbenzhiltonhead.com/what-does-the-coffee-cup-mean-on-the-mercedes-benz-instrument-cluster/

[202] GLESS supra note 157 at 204.

[203] Translated from the original: "São admissíveis as provas que não forem proibidas por lei".

[204] This non-restrictive rule also applies to evidence gathering methods from articles 171º-189º PCCP, nevertheless they aren't mentioned as they aren't the focus of this analysis. About this subject *see*

The law envisages an *openness* of the evidentiary legal framework that has two main effects: the admissibility of new types of evidence different from the ones already regulated by law (atypical evidence)[205] and that any fact may be proven by any type of evidence.[206] Thus, advanced safety systems's findings could –at least at first sight– be admitted as new type of evidence.

Nevertheless, according Paulo Sousa Mendes, Pinto de Albuquerque, and most Portuguese scholarship, this "freedom" in accommodating any atypical evidence is illusory, as from the start their admissibility is subject to constitutional and legal limits inherent to the accusatory structure of the Portuguese criminal proceeding.[207] This naturally means that the first limit to the "free" admissibility of new evidence regardless of its previous regulation resides in the respect for fundamental rights, freedoms and guarantees enshrined in the Portuguese Constitution.[208] Portuguese scholarship makes a distinction between *absolute prohibitions*, under the exemplificative list from art.126, nº 2 PCCP, which refers to evidence that is obtained through the violation of fundamental rights, and that are obtained in breach of human dignity (art. 1º PCR), physical and moral integrity (art. 25º PCR), and *relative prohibitions* obtained under abusive intrusion in private life is admissible. The latter are prohibited when obtained in breach of the right to privacy without proper consent.[209]

---

ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem",* 2018, p. 332 *and* MARQUES DA SILVA, Germano - *"Curso de Processo Penal II"*, 2008, p. 136, 137.

[205] MARQUES DA SILVA, Germano - *"Curso de Processo Penal II"*, 2008, p. 136, 137.

[206] ALBERGARIA, Pedro Soares, et al. – "*Comentário Judiciário do Código de Processo Penal"*. t.II. (art. 125º), 2019, p. 29, 30. *Also refer* to ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem",* 2018, p. 332. *And* SILVA, Sandra Oliveira – *"Legalidade da Prova e Provas Proibidas" in* Revista Portuguesa de Ciência Criminal, n.º 4, 2011, p. 13: "(…) em princípio, a liberdade de escolher indiferentemente qualquer dessas fontes tipificadas de conhecimento, seja qual for a natureza dos factualidade a provar."

[207] SOUSA, Paulo Mendes – *"Lições de Direito Processual Penal",* 2014, p. 173; ALBUQUERQUE supra note 207 at 332.

[208] ROBALO, Inês –*" Verdade e Liberdade – A Atipicidade da Prova em Processo Penal",* 2012, p. 49. Available at: https://repositorio.ucp.pt/bitstream/10400.14/15696/1/Verdade%20e%20Liberdade%20-%20A%20Atipicidade%20da%20Prova%20em%20Processo%20Penal%20-%20In%C3%AAs%20Robalo.pdf

[209] SILVA, Sandra Oliveira – *"Legalidade da Prova e Provas Proibidas" in* Revista Portuguesa de Ciência Criminal, n.º 4, 2011, p. 30, 32, 33. See also: ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem",* 2018, p. 335-337; ALBERGARIA, Pedro Soares, et al. – "*Comentário Judiciário do Código de Processo Penal"*. t.II. (art. 126º), 2019, p. 39 – 64.

This means any evidence that is feasible to restrain the right to privacy to some degree is subject to the limit of the rule of law (*reserva de lei*), and only admitted to the strictly necessary to safeguard other colliding rights and interests safeguarded by the constitution, in light of the principle of proportionality, as enshrined in art. 18º, n. º 2 PCR and art. 34º, nº 2, 4 PCR.

The third limit to the free admissibility of evidence resides in the respect for the already typified frameworks for each means of evidence, to prevent the non-restriction rule to become a gateway for fraudulent new types of evidence used to replace already regulated evidence. In fact, this matter was already discussed by the Portuguese Supreme and Constitutional Courts regarding the replacement of identification evidence *(prova por reconhecimento)*[210] by a testimony during trial. [211] Before PCCP's legislative reform in 2007, it used to be a practice in Portuguese courts to admit testimonial evidence to replace the typical recognition evidence - arts. 147º and 148º PCCP- under an atypical "identification" performed by the witness during the trial, as it was interpreted that the recognition's rules only applied to pre-trial stages. Now it is clear due to the changes made in 2007 in the legal provision for recognition evidence– art. 147º/7 PCCP - that the procedure envisaged to the evidence by recognition is prescriptive and must be observed in all stages of the criminal proceeding, including the trial hearing. Each typical evidence was previously reasoned, designed and regulated according to its epistemological background to attain a reliable and concrete cognoscitive result whilst assuring the respect for the defendant's fundamental rights. Hence, the use of any typical or atypical evidence in replacement of an already regulated typical evidence to distort their rules might originate in an unreliable and defective cognitive result.[212]

[210] About this specific type of evidence see generally: DUARTE, Eurico Balbino – *"Making Of – A Reconstituição do Facto no Processo Penal Português"* in Prova Criminal e Direito de Defesa: Estudos sobre a Teoria da Prova e Garantias da Defesa em Processo Penal, 2013, pp. 7-24;

[211] See Supreme Court Ruling: Ac.STJ, 20-09-2017. 1353/13.6GBABF.E1.S1 available at: http://www.dgsi.pt/jstj.nsf/-/AEAC8214F656C2FC802582560040210A and Constitutional Court Ruling: AC. TC nº 137/2001. 78/2000 available at: https://dre.pt/dre/detalhe/acordao/137-2001-2875168 each court recognized that the disrespect for the prescriptive steps regulated for recognition evidence at any stage of the proceeding was illegal as it harmed the fundamental right of defence. By violating the procedure for an adequate recognition, the defendant wasn't able to refute adequately that he/she wasn't identified correctly as there was no opportunity to compare her/him to any other individuals.

[212] ALBERGARIA *supra note* 207 at 32. Refer also to: DÁ MESQUITA, Paulo, et al. - "*Comentário Judiciário do Código de Processo Penal"*. t.II. (art. 147º), 2019, pp. 350-354, 356-357.

Taking this into consideration, Portuguese scholarship - Medina de Seiça, Germano Marques da Silva, Sandra Silva Olivera - have been conclusive that the non-restriction rule of art. 125º PCCP must not be confused with unrestricted fungibility of means of evidence as it might affect the right to a proper defence. [213] This understanding is also shared by the Italian law through the principle of non-replacement of means of evidence *(principio di non sostituibilità)*.[214]

Besides this, a new type of atypical evidence must be subjected to a final requirement: it must be adequate and suitable to assess the facts it is meant to prove, as different from the already regulated evidence it was not previously reasoned. It must be ensured that the evidence to be admitted is suitable for producing a reliable and useful cognitive result.[215]

To validate the admissibility of AI-generated evidence, the first step is to assert if it complies with the protection of fundamental rights as AI systems, due to their characteristics entail risks for fundamental rights, mainly the right to a fair trial, to contradictory, impartiality, and respect for privacy and data protection.[216]

### 4.1.1 The Black Box Paradox

The poignant distinctive trait of AI machine evidence from any other evidence resides on its autonomy. AI systems render their own assertions and convey messages as a result from the analysis they make from their surrounding environment. Considering our leading example: in the course of an investigation a "car" will be able to convey if the driver was drowsy or distracted before an accident has occurred.[217]

The problem that arises from using self-learning algorithms for truth assessment in criminal justice is that the algorithmic evidence results are inherent to *inscrutability* and *opacity*.[218]

---

[213] SILVA, Sandra Oliveira – *"Legalidade da Prova e Provas Proibidas" in* Revista Portuguesa de Ciência Criminal, n.º 4, 2011, p. 19; 199; SEIÇA, Alberto Medina de - *"Legalidade da prova e Reconhecimentos «atípicos» em processo penal: notas à margem de jurisprudência (quase) constante"* in Liber Discipulorum para Jorge de Figueiredo Dias, organizado por Manuel da Costa Andrade, 2003, p. 1412, 1413.
[214] ALBERGARIA *supra note* 220 at 33.
[215] *Id* at 32.
[216] COM (2020) 65 final. Brussels, 19.02.2020 – *White Paper on Artificial Intelligence – A European approach to excellence and trust*, p. 10.
[217] About the functioning of these systems please refer to section 4.
[218] MSI-AUT(2018)05. Committee of Experts on Human Rights Dimensions of Automated Data Processing and different forms of Artificial Intelligence - *A study of the implications of advanced digital technologies*

Unlike the early forms of AI which relied on expert systems and rule-based reasoning where the machine did restrictively what it was told to do, contemporary learning systems are more complex. The more layers of data are added and analyzed by the algorithm, the more difficult it is to trace its underlying logic.[219] Although the degree of algorithm complexity may vary according to the autonomy level of each AI system and its functioning, resulting in more accessible, and therefore more scrutable decision-making processes[220] (which is not the case of AI monitoring systems in automated vehicles as they rely on ANNs combined with other machine learning algorithms such as SVMs)[221], it does not imply open access to the algorithm as the latter is usually subject to intellectual property rights. In other words, even if the complexity of the algorithm is at a lower level allowing an explanation of its functioning, the access to the algorithm is not always granted as most of these systems are embedded in consumer products developed by private companies.[222]

According to Burrel, an algorithm "is opaque in the sense that if one is a recipient of the output of the algorithm, rarely does one have any concrete sense of how or why a particular classification has been arrived at from inputs".[223] The author distinguishes three categories of opacity: i) Opacity as intentional corporate or state secrecy, this is the type of opacity found in consumer products, such as the AI monitoring systems under analysis; this is an intentional form of self-protection by corporations to maintain their trade secrets and competitive advantage; ii) Opacity as technical illiteracy, the design and operation

*(including AI systems) for the concept of responsibility within a human rights framework,* 2018, p. 4, 14-16;

[219] VEALE, Michael – *"Algorithms in the Criminal Justice System. A Report by Law Society of England and Wales",* 2019, p. 4 *see also* c2018, p. 15. *And* NUTTER, Patrick – *Machine Learning Evidence: Admissibility and Weight*. Journal of Constitutional Law, vol.21:3, 2019, p. 928.

[220] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 204.

[221] In order to better understand the underlying functioning of AI monitoring systems in automated vehicles please refer to 4; *Also* MSI-AUT(2018)05. Committee of Experts on Human Rights Dimensions of Automated Data Processing and different forms of Artificial Intelligence - *A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework,* 2018, p.16: "While some forms of learning systems enable the underlying logic to be traced and understood (for example, those which utilise decision-trees), others (including those that utilise neural networks and back propagation) do not."

[222] MSI-AUT(2018)05. Committee of Experts on Human Rights Dimensions of Automated Data Processing and different forms of Artificial Intelligence – "*A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework",* 2018, p. 16

[223] BURREL, Jenna – *"How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms"* in Big Data and Society (I-12), 2016, p. 1.

mode of learning systems requires specialized skills and knowledge on software engineering, which common citizens aren't expected to have; iii) Opacity as the way algorithms operate at the scale of application, this resumes to the above mentioned inscrutability inherent to the own functioning of the learning algorithms. Efficient ML algorithms "possess a degree of unavoidable complexity" Many algorithms are "multi-component systems", besides the "number of lines and pages of code, the number of different programmers in the engineering team, and the multitude of interlinkages between modules and sub-routines creates challenges of scale and complexity that are distinctive to machine learning models".[224]

It is the combination of these two factors, *inscrutability* and *opacity* that results in the *black box paradox*, resuming the difficulty for general users and even experts to understand and explain the machine's outputs.[225] Resuming our example: the camera or the component of the drowsiness detection system in the car is visible to the driver, however, (s)he does not know what is behind the evaluative process, and experts as well might have difficulty understanding the reasoning behind the decision process or not even have access to the full code.[226] This problem is even bigger when the machine provides an unexpected or incorrect result (e.g. FAIR negotiation bots).[227]

---

[224] *Id*. at 3-5.

[225] MSI-AUT(2018)05. Committee of Experts on Human Rights Dimensions of Automated Data Processing and different forms of Artificial Intelligence - *A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework.* 2018, p. 16 Regarding the introduction of the black box paradox in portuguese criminal justice see: RODRIGUES, Anabela Miranda – *"Inteligência Artificial no Direito Penal – A Justiça Preditiva entre a Americanização e a Europeização"* in A Inteligência Artificial no Direito Penal, 2020, p. 25. and SOUSA, Susana Aires de – *"Não fui eu, foi a máquina: Teoria do Crime, Responsabilidade e Inteligência Artificial"* in Inteligência Artificial no Direito Penal, 2020, p.67

[226] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials"*. 2020, p. 211

[227] MANSON, Stephen; SENG, Daniel – *"Artificial Intelligence and Evidence"*, 2021, p. 245.
A popular example of how machine learning may produce unexpected results even to its programmers was Facebook AI Research (FAIR), a project developed by Facebook company (now named META) in 2017. FAIR was supposed to be a negotiation program. The programmers were training two chatbots to develop a negotiation language, however, the chatbots ended up developing a different language for which they were not initially programmed. As they were not providing the expected outcome the engineers turned off the simulation. *See* FAIRS' project: LEWIS, Mike; YARATS, Denis; BATRA, Dhruv – *"Deal or no Deal? Training AI Bots to Negotiate"*. [Online].2017.[Last access: 17.05.2022]. Available at: https://engineering.fb.com/2017/06/14/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/ *see also* KUCERA, Roman – *"The Truth Behind Facebook AI Inventing a New Language"*.[Online].2017.[Last access:17.05.2022]. Available at: https://towardsdatascience.com/the-truth-behind-facebook-ai-inventing-a-new-language-37c5d680e5a7

From the perspective of criminal justice, it is necessary to analyze if the opacity resulting from the AI-generated evidence may affect the defendant's rights.

The defendant status in the context of a criminal procedure implies the attribution of a set of safeguards that must be respected during the proceeding, known as the defence rights. These rights are enshrined in art. 32º of the Portuguese Constitution and art. 61º PCCP. The respect for the defendant's rights and criminal safeguards are also part of the nuclear values of the European Union (EU) as corollaries of fundamental rights, as provisioned in arts.48º-50º of the European Charter of Fundamental Rights, and together they constitute the core of the right to a fair trial reflected in art.47º of the European Charter of Fundamental Rights, art. 6º of the European Convention of Human Rights.

The admission of machine evidence might constitute a risk to the defence rights in the dimension of the principle of contradictory *(princípio do contraditório)*, presented in art. 32º/5º *in fine* CPPC. The principle of contradictory is inherent to the accusatory structure of the Portuguese criminal procedure (art. 32º/5 PCR) and determines that the trial hearing and all inquiry acts *(atos instrutórios)* must be submitted to contradictory, this includes all evidence produced during the trial according to arts. 340º CPPC and 327º/2º CPPC. According to Germano Marques da Silva, this means that all criminal prosecution must take into consideration both accusation and defence motivations.[228] The principle of contradictory guarantees the defendant[229] the right to participate effectively in the proceeding, including, *inter alia*, the right to be present, hear, follow the proceeding, and to challenge, effectively and efficiently, the evidence produced against him.[230] Likewise, Germano Marques da Silva, this implies the right to an effective "control" of the produced evidence, including being able to challenge its relevance and reliability.[231] In this case, for the defence right to be ensured it would be necessary for the defendant to be able to challenge and examine the result behind the drowsiness and distraction detection system,

[228] FIDALGO, Sónia – *"A Utilização de Inteligência Artificial no Âmbito da Prova Digital – Direitos Fundamentais (ainda mais) em Perigo"* in Inteligência Artificial no Direito Penal, 2020, p. 144.
[229] The principle of contradictory *lato sensu* is extensive to all proceeding subjects, nevertheless, this analysis is centered on the defence rights perspective.
[230] ZAVRŠNIK, Aleš – *"Criminal Justice, Artificial Intelligence Systems, and Human Rights"* in ERA Forum 20. 2020, p.576 : " In order to ensure effective participation in a trial, the defendant must also be able to challenge the algorithmic score that is the basis of his or her conviction."
[231] MARQUES DA SILVA, Germano - *"Curso de Processo Penal II"*, 2008, p. 192.

which will not be possible if full information regarding the functioning of the algorithm is not disclosed.[232]

The black box problem is also raising new concerns regarding the principle of equality of arms *(princípio da igualdade de armas)* in the context of criminal proceedings evolving machine evidence, as this new type of evidence is feasible to generate significant knowledge impairment.[233] Similarly to the principle of contradictory, the principle of equality of arms is inherent to the accusatorial structure of the criminal procedure, and a dimension of the fundamental right to a fair trial. Although there is no explicit reference to this right in legal provisions of ECHR, nor in the portuguese criminal procedural code and portuguese constitution provisions, its legal basis has been crafted by scholarship and jurisprudence from art.º 6.º, nº 1 ECHR, and from the Portuguese legal perspective from art.º 20º, nº 4 PCR.[234]

When alluding to this principle the unbalanced nature between the parties in criminal proceedings must be considered. It cannot be expected full equality due to the opposition between public prosecution and individuals, mainly during the investigation when the prosecution has an insurmountable advantage regarding the applied investigation methods combined with the possibility to adopt cautionary measures.[235] Besides, this imbalance does not refer exclusively to the prosecution side, as the defendant benefits from a set of fundamental warranties such as the right to remain silent, the right not to incriminate him/herself, the right to present a closing statement (artº 361º PCCP) , and the in *dubio pro reo* principle, which would be harmed if the principle of equality of arms was interpreted at its full extension.[236] That being said, the equality of arms must be

[232] VEALE *supra note* 215 at 57.

[233] RODRIGUES, Anabela Miranda – *"Inteligência Artificial no Direito Penal – A Justiça Preditiva entre a Americanização e a Europeização"*. in A Inteligência Artificial no Direito Penal, 2020, p. 16.

[234] See PAGALLO, Ugo; QUATTROCOLO, Serena – *"The Impact of AI in Criminal Law, and its Twofold Procedures"* in Research Handbook on the Law of Artificial Intelligence, 2018, p. 396. *And amongst Portuguese scholarship:* MARQUES da SILVA, Germano – *"Curso de Processo Penal I",* 2010, p. 78, 79; MOREIRA, Vital; CANOTILHO, Gomes – *"Constituição da República Anotada" ( artº 20)*, vol. 1, 2007. *See also:* Constitutional Court ruling: AC.TC.160/2010, n.º 834/09, 27.04.2010 available at: http://www.tribunalconstitucional.pt/tc/acordaos/20100160.html

[235] LIGETI, Katalin – *"Artificial Intelligence and Criminal Justice."* in AIDP-IAPL International Congress of Penal Law. 2019, p. 11. And MARQUES da SILVA, Germano – *"Curso de Processo Penal I",* 2012, p. 78, 79.

[236] D*IAS,* Jorge Figueiredo – *"Sobre os Sujeitos Processuais no Novo Código de Processo Penal"* in Jornadas de Direito Processual Penal: O novo código de processo penal, 1998, p. 29 :"*Este principio (…) não pode, sob pena de erro crasso, ser entendido como obrigando ao estabelecimento de uma igualdade matemática ou sequer lógica. Fosse assim e teriam de ser fustigadas pela crítica numerosas normas com bom*

interpreted in accordance with the logical and material structure of accusation and defence, and its dialectics,[237] meaning that each party must be given the same reasonable opportunity to present their case under conditions that do not place them at a substantial disadvantage over the opponent.[238] According to Quatrocollo and Pagallo, the use of inculpatory evidence that relies exclusively on algorithm processes such as neural networks of AI systems has the potential to generate a "huge disproportion between the parties to a criminal proceeding" as the defendant has almost no opportunity to challenge the evidence produced against him.[239]

Another fundamental right that might be harmed with the use of AI systems in the context of a criminal proceeding is the right to respect for private and family life enshrined in art.8º ECHR. AI systems as a general purpose technology may pervade and cause impact in the entire fabric of society. One of the most important societal impacts of AI is the intrusion in private areas of our lives.[240] Besides the fact that AI systems themselves rely on data to function, the number of AI applications that work by processing and storing biometric data, performing facial recognition, and surveilling human activities is increasing. These systems may find their way to justice enforcement as strong means of evidence. AI-driven surveillance evolves systems that are able of capturing images (personal image), storage and process personal biometric data, affecting our general privacy, identity, and autonomy, potentially creating a constantly watched, followed and identifiable environment.[241]

---

*fundamento (…) como os da inviolabilidade do direito de defesa, da presunção de inocência do arguido, ou do in dúbio pro reu".*

[237] *Id* at 30.

[238] LIGETI supra note 230 at. 396. *see also* Guide on Article 6 of the European Convention on Human Rights (criminal limb). 2019, p.132. and MARQUES DA SILVA, Germano – *"Curso de Processo Penal"*. Vol I. 2010, pp.78-79. Also refer to the Portuguese Supreme Court of Appeal ruling referring this subject: AC. STJ, n.º 251/15.3GDCTX.L2.S1, 07.03.2018 available at: http://www.dgsi.pt/jstj.nsf/954f0ce6ad9dd8b980256b5f003fa814/d4dd16b72a700f83802582c7004a9777?OpenDocument

[239] PAGALLO, Ugo; QUATTROCOLLO, Serena – *The Impact of AI on Criminal Law and its Twofolds Procedures* in Research Handbook on the Law of Artificial Intelligence, 2018, pp. 395, 396. *See also* ZAVRŠNIK, Aleš – *"Criminal Justice, Artificial Intelligence Systems, and Human Rights"* in ERA Forum 20. 2020, p. 577 : At least some degree of disclosure is necessary in order to ensure a defendant has the opportunity to challenge the evidence against him or her(…)"

[240] BEN-ISRAEL, Issac, et al. – *"Towards Regulation of AI Systems- Global Perspectives on the development of a legal framework on Artificial Intelligence systems based on the Council of Europe's standards on human rights, democracy and the rule of law",* 2020, p. 29.

[241] *Id* at 30.

According to art. 126º/3 PCCP unless in specific cases predicted by law, all evidence that intrudes private life, home, correspondence, and telecommunications, without the previous consent is null. Meaning all evidence that implies an abusive intrusion in the defendants (or others) rights to private life as enshrined in arts. 26.º, n. º 1 PCR, 32.º, n.º 8º PCR, 34. º PCR are prohibited.[242] The Portuguese scholarship refers to these cases as relative prohibitions *(proibições relativas de prova)*.[243]This means the right to private and familiar life may be curtailed when predicted by law or waved by its titular, however it must be submitted to the rule of law *(reserva de lei)* and to the principle of proportionality (art. 18º PCR; 34º/2/4 CRP) in a way they must affect the defendant's rights to the bare minimum in proportion to the colliding right.[244]

When it comes to advanced driver drowsiness and attention detection systems, the principal concern resides in the assessment of its risks to the users' privacy and data protection (art. 8º ECFR). Advanced safety systems' operation implies monitoring human behaviour through cameras and sensors, collecting and processing data in real-time, which may lead to the conclusion that rights to privacy and data protection might be at stake. This seems to be the reason why such systems are initially considered high-risk in accordance with art. 6º and Annex II from the AI Act proposal, and thus subject to its technical and conformity assessment requisites.

Nevertheless, the AI Act proposal seems to contradict itself, as in art. 2º, n.º 2 it excludes the safety systems from Regulation (EU) 2019/2144 from the vast majority of its application scope.[245] There is no clear reason for this discrepancy which may be, according to Bomhard and Merkle, justified as an editorial error that may be corrected with the deleting of Annex II or art. 2º, n.º 2.[246] However, deepening this matter, there

[242] SILVA, Sandra Oliveira – *"Legalidade da Prova e Provas Proibidas"* in Revista Portuguesa de Ciência Criminal. Nº4, 2011, p.33

[243] See TRIUNFANTE, Luís Lemos, et al. – "*Comentário Judiciário do Código de Processo Penal"*. t.II. (art. 128º), 2019, p.53-65. And SILVA supra note 256.

[244] SILVA, Sandra Oliveira – *"Legalidade da Prova e Provas Proibidas"* in Revista Portuguesa de Ciência Criminal, n. º 4, 2011, p.34

[245] COM (2021) 206 final. Brussels, 21.04.2021 - Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts (Artificial Intelligence Act), articles 6º, 2º/2 and 84º.

[246] BOMHARD, David; MERKLE, Marieke – *"Regulation of Artificial Intelligence. The EU Commission's Proposal of an AI Act"* in EuCML, Issue 6, 2021, p. 260.

might be a reason behind the choice of treating these specific safety systems differently from other high-risk systems.

The main reason might relate with consistency concerns between the AI Act and the already existing sectorial product safety legislation. The explanatory memorandum of the AI Act is explicit regarding consistency matters, to avoid duplication of conformity requirements and minimize additional burdens. In fact, it excludes its direct application to aviation and car products covered by relevant Old Approach Legislation,[247] and refers that consistency assessment regarding products already regulated under safety legislation will follow third party assessment procedures as already established under each relevant sectoral product safety legislation.[248] This seems to be reflected in the AI Act art. 2º, n.º 2 as all the exceptions consist of regulations and directives already referring to type-approval, and conformity assessment for motor-vehicle, aviation and maritime equipment (e.g., Directive 2014/90/EU, Regulation (EC) No 300/2008, Regulation (UE) 2019/2144).

It is important to take into consideration that to the specific sector of vehicle regulation is applicable the legal framework of UNECE World Forum for Harmonization of Vehicle Regulations (WP.29). Therefore, harmonized rules and technical requirements for automated vehicle systems, should be adopted and promoted at international level in UNECE's World Forum for Harmonization of Vehicle Regulations (WP.29). This means that UN regulations or other regulatory acts adopted under WP.29 regarding the software update and conformity of such systems should be applied.[249]

Notwithstanding, such systems remain under the periodic evaluation and review obligation of art. 84º, probably to assess if substantial modifications in the AI systems justify additional or new *ex ante* re-assessments, implying changes in the way the AI Act approaches such systems.

The Regulation (UE) 2019/2144 on the type-approval requirements for motor vehicles, their trailers, systems, components, and technical units takes into consideration the aforementioned risks regarding privacy and data protection issues. In recital 10, it is set that driver drowsiness and attention warning systems should function without biometric

[247] COM (2021) 206 final. Brussels, 21.04.2021 - Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts (Artificial Intelligence Act), p. 4.
[248] *Id* at 14.
[249] Refer to Recitals (23), (27) Regulation (EU) 2019/2144 of 27 November 2019, p. 5.

information of drivers or passengers, including facial recognition. It also states that all technological progress regarding these systems must be considered in every evaluation of the existing legislation to ensure its strict adherence to privacy and data protection rules.[250]

Every processing of personal data about the driver or information regarding the driver drowsiness and attention should be carried in accordance with GDPR. Furthermore, those systems must be designed in ways that "do not continuously record nor retain any data other than what is necessary in relation to the purposes for which they were collected or otherwise processed within the closed-loop system".[251] This requirement is materialized in art. 6º, n.º 3 of the regulation, with the additional safeguard that the collected data "shall not be accessible or made available to third parties at any time and shall be immediately deleted after processing".[252]

To ensure that these systems are placed on the market complying with the requirements in the regulation, manufacturers shall demonstrate that those are type-approved in accordance with the requirements presented in the regulation, according to art. 4º.

Also, these systems will be submitted to periodic review and reporting obligations. As predicted in art. 14º, "the Commission shall investigate whether those safety measures and systems act as intended" and submit on an annual basis a report to the European Parliament and Council on the activities of UNECE's World Forum for Harmonization of Vehicle Regulations (WP.29) with regards to the progress made in on the requirements set out in Articles 5 to 11.

These systems' goal is to function as mere safety tools that convey a message to trigger a reaction that envisages avoiding a road accident, not as intrusive surveilling systems *per se*. In fact, if driver drowsiness and attention warning systems' functioning comply with the regulation requirements their comprised risk to privacy and data protection will be substantially low. However, one must consider if the technical reality of such systems as described in section 4., will in practice be fully compatible with non-intrusive requirements, avoiding falling in a presumption of regularity.[253] Some solutions regarding

[250] Refer to Recital (10) Regulation (EU) 2019/2144 of 27 November 2019, p. 2.
[251] Refer to Recital (14) of Regulation (EU) 2019/2144 of 27 November 2019, p. 3.
[252] *Id* at art. 6º, n. º 3.
[253] GLESS supra note at 250.

the design of non-intrusive drowsiness systems are already in place and revealed their effectiveness was equivalent to intrusive techniques.[254] Nevertheless, even if there is the possibility of these systems to not materialize substantial harm, they still have for its nature a potential to be harmful.

### 4.1.2 Bias by design and automation bias

The black box behind the learning algorithm's functioning may also lead to the harming of the right to an impartial ruling as an element of the right to a fair trial, and impact adversely on the principles of presumption of innocence enshrined in art. 6º, n.º2 ECHR, art. 48º ECFR and art. 32º, nº 2 PCR, generically referring that "everyone charged with a criminal offence shall be presumed innocent until proved guilty according to law".

Nonetheless, scholarship and jurisprudence have been entailing important corollaries from this principle,[255] the "burden of proof on the accuser"[256]/prosecution, meaning the defendant should have to prove his/her innocence and a "standard of proof beyond reasonable doubt".[257] Likewise, Figueiredo Dias,[258] infers that the presumption of innocence safeguards the defendant from a burden of proof, which as mentioned by Germano Marques da Silva,[259] should fall over the prosecution that must provide evidence that reflects without reasonable doubt that the defendant committed the crime. This leads to an important mediating, the principle of *in dubio pro re*o, any "*non liquet*" situation in regard to production of evidence must act in favour of the defendant.[260]

Other corollaries, such *as the right on the judge not to start with the preconceived idea that the accused is guilty* are also mentioned by jurisprudence, amongst Portuguese scholarship Germano Marques da Silva mentions that the defendant should not be forced to prove his/her innocence to exclude a *presumption of previous guilt*, which means the judge should *ab initio* assume in any circumstances that the defendant is innocent so the

[254] See generally OLIVEIRA, Licínio – "Driver Drowsiness Detection Using Non-Intrusive Signal Acquisition", 2018. Available at: https://repositorio-aberto.up.pt/bitstream/10216/113802/2/276791.pdf
[255] *SACHOULIDOU, Athina – "OK Google: is (she) guilty?"* in Journal of Contemporary European Studies, 2021, p. 4.
[256] *Id.*
[257] *Id.*
[258] DIAS, Jorge Figueiredo – *"Direito Processual Penal"*, 1974, p. 122.
[259] SILVA, Germano Marques da – *"Curso de Processo Penal", 2008, p.* 84.
[260] Id at 84.

defendant does not have to reunite efforts to prove his/her innocence.[261] Such principles might be potentially harmed when the court is confronted with AI generated evidence, as its associated efficiency may generate in the judge an automation bias, this phenomenon will be approached in the next lines along with the analysis regarding the judge's impartiality.

According to the provision of art. 6º/1 ECHR the right to a fair trial requires a tribunal to be impartial, in the sense it must decide in the absence of prejudice or bias.[262]

Despite the natural human ease in relying on machine decisions, it is important to take into consideration that AI systems may generate fragile evidence, as ML techniques may learn from biased or poor quality data, and therefore produce an incorrect and/or biased decision.[263] As observed by the European Commission for the Efficiency of Justice (CEPEJ) the neutrality of the algorithm is a myth, the design of algorithms is made by humans who may consciously or unintentionality influence algorithms' predictions and estimates by transferring their own assertions and valuations to the decision models.[264]

Therefore, there is a risk that a court's decision supported by AI-generated evidence might be relying on assertions that result from hidden biases, harming the right to an impartial ruling and the principles of equality and non-discrimination (art. 14º ECHR and 13º PCR).

Taking the example of drowsiness systems, their output may be imprecise or ambiguous, as it may include biased algorithms or biased standardized data that affect the output, mainly when referring to consumer products that are likely to have hidden patterns of subjectivities (eg., Volkswagen and Uber).[265] The choice of a particular design to capture

[261] SILVA, Germano Marques da – *"Curso de Processo Penal",* 2008, p. 84.
[262] Guide on Article 6 of the European Convention on Human Rights (criminal limb), 2019, p.25.
[263] FIDALGO, Sónia – *A Utilização de Inteligência Artificial no Âmbito da Prova Digital – Direitos Fundamentais (ainda mais) em perigo* in A Inteligência Artificial no Direito Penal. 2020, p. 145 see also MANSON, Stephen; SENG, Daniel – *"Artificial Intelligence and Evidence"*,2021, p. 245-247.
[264] European Commission for the Efficiency of Justice (CEPEJ) – *"European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment"*.2018, p.57; VEALE *supra note* 215 at 18: "Algorithms encode assumptions and systematic patterns which can result in discriminatory outputs or downstream effects. The way data used as input to systems is labelled, measured and classified is subjective and can be a source of bias."
[265] CHESTERMAN, Simon – *"Trough a Glass, Darkly: Artificial Intelligence and The Problem of Opacity"* in NUS Law Working Paper nº 2020/1, 2020, p. 8: "Volkswagen, for example, wrote code that gamed tests used by regulators to give the false impression that vehicle emissions were lower than in normal usage. Uber similarly designed a version of its app that identified users whose behavior suggested that they were working for regulators in order to limit their ability to gather evidence."

a driver's face or body position and the sample population used might have a serious impact on the drowsiness assertion. For instance, if the dataset used as sample to train the algorithm does not consider variations in eyelid positioning across ethnicities may generate an erroneous output for interpreting individual variations as a sign of sleepiness[266] or even the fact if the system is apt to evaluate with accuracy individuals wearing glasses.

Bias can be learned at least in two ways: when the learning process is affected by biased data, because they were unintentionally introduced by the programmers or when programmers' prejudice intentionally affects the data used to train the algorithms (e.g., cases of software design favoring the corporate self-interest).[267] These cases are defined by Roth as *falsehood by human design*, in this case the machine conveyance might be false or misleading as it was programmed to render inaccurate information.[268] The second situation occurs when unintended biases result from the learning process through the weighting of variables collected from a sample population; in this case, the learning process itself draws biased inferences[269] (e.g. Tay Tweets and Amazon's recruiting tool).[270] According to Roth, these situations consist of *falsehood by machine-learning design*, and are a result of the complex nature of machine learning itself and its interaction with the world.[271] This is one poignant difference between the machine and human

[266] GLESS supra note at 217.
[267] *Id at* 206;
[268] ROTH, Andrea – "*Machine Testimony*",2017. Yale Law Journal, Vol. 126. Nº.1,p. 1991.
[269] CHESTERMAN, Simon – "*Trough a Glass, Darkly: Artificial Intelligence and The Problem of Opacity*" in NUS Law Working Paper nº 2020/1, 2020, p. 12;
[270] *Id :* "Amazon's résumé screening algorithm, which was trained on ten years of data but had to be shut down when programmers discovered that it had 'learned' that women's applications were to be regarded less favorably than men's".
Tay Tweets was a chatbot developed by Microsoft in 2016, which objective was to engage the teenager audience to conduct research on natural language processing (a branch of AI that allows systems to interpret human speech and text). Tay was launched on the social network platform Twitter, and it took less than 24 hours to start displaying racist and inflammatory hate speech. Despite the implemented filtering, Tay ended up learning from internet "trolls" and Microsoft had to deactivate it a short time after for adjustments. *See* LEE, Peter – *"Learning from Tay's introduction"*. [Online]. 2016. [Last access: 17.05.2022.] Available at: https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/ *see also* HUNT, Elen – *"Tay, Microsoft's AI Chatbot, Gets Crash Course in Racism from Twitter"*. [Online].2016. [Last access: 17.05.2022.] Available at: https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter
[271] ROTH, Andrea – "*Machine Testimony*". 2017. Yale Law Journal, Vol. 126. Nº.1,p. 1991

reasoning, while machines make correlations from any biased data and provide mechanical decisions, humans "fall back on common sense"[272] to weigh their decisions.

The lack of impartiality and harm of the presumption of innocence, mainly on its dimension of *in dubio pro reo*, may result as well from a phenomenon named *automation bias*, also know as *presumption of reliability*.[273] This effect already generated its share of concern amongst the admissibility of digital evidence, meaning there is a general complacency from judges in considering informatic systems flawless and reliable.[274] There is a "general belief in the superior judgement of automated aids"[275] that may generate the assumption that machines' assertions are correct and true. This may difficult a proper analysis of the underlying functioning of the system that generated the machine "testimony" before being admitted in court, generating to the defendant an excessive burden to challenge the machine evidence and face the risk to be proven guilty based on non-contested automated evidence.

### 4.1.3 Right to Explanation, Reliability Testing

There are other dangers that derive from the black box problem, aside from bias by design and automation bias, such as machine inarticulateness. In these cases, the machine is imprecise, ambiguous or experiences a breakdown in its reporting capacity, due to human design choices, operation errors and machine's degradation.[276]

A solution to enhance accountability and transparency may reside in a right to explanation, in order to humans to achieve sufficient knowledge on how the machine generated its outputs. Scholars, such as Goodman and Flaxman[277], Selbst and Powles[278] argue that such a right derives from articles 13º-15º in connection with Recital 71 of the General Data Protection Regulation (GDPR), which must consist in the right of

[272] MANSON, Stephen; SENG, Daniel – *"Artificial Intelligence and Evidence"*,2021, p. 246.
[273] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 206. *And,* MANSON, Stephen; SENG, Daniel – *"Artificial Intelligence and Evidence"*, 2021, p. 246.
[274] *Id at* 248. Among Portuguese scholarship *see* also RAMALHO, David Silva – *"Métodos Ocultos de Investigação Criminal em Ambiente Digital"*, 2017, p. 259.
[275] ZAVRŠNIK *supra note* 142 at 13.
[276] ROTH, Andrea – "*Machine Testimony"*. 2017. Yale Law Journal, Vol. 126. Nº.1,p. 1993-2000
[277] See generally GOODMAN, Bryce; FLAXMAN, Seth – *"EU Regulations on Algorithmic Decision Making and a Right to Explanation"* in AI Magazine. Vol.38, nº3 and
[278] See SELBST, Andrew; POWLES, Julia – "Meaningful Information and the Right to Explanation" in International Data Privacy Law. Vol.7. nº4.2017

"meaningful information about the logic involved" in automated decisions. Such a right to an explanation should be interpreted functionally, flexibly and should, at a minimum, enable a data subject to exercise their rights under the GDPR and human rights laws.

A right to explanation should be granted when they are used as evidence in the context of a criminal trial, in order to provide the necessary transparency on its functioning for the defendant to challenge effectively the evidence produced, and for the judge to form a sustainable conviction.

A necessary step before admitting a new type of evidence in the criminal proceeding is to attest its relevance and reliability. This criteria is found in art. 340º/4, a), b) PCCP. The use of evidence is rejected when the evidence is not relevant, nor reliable, which means when it has no particular connection with the fact it is offered to prove, and it is inadequate when it is not reliable according to current scientific methods and knowledge.[279] According to Sandra Oliveira e Silva in order to evaluate the reliability of new types of evidence, the judge acts as a "gatekeeper", assuming effective control of the evidence's scientific reliability, by adopting the necessary methodological tools.[280] Some scholars,[281] refer to the criteria that has been usually performed, mainly in the USA, under the *Daubert* test, as an example that should be adapted to the standards of the Portuguese juridic system.[282] According to Nutter, it is likely that substantive evidence generated by ML techniques will be used in court under the form of expert testimony and being subject to Daubert test criteria: "whether the theory or technique can be or has been tested; whether the theory or technique has been subject to peer-review publication; third, the existence of error rates; and fourth whether the theory or technique enjoys general acceptance in the field or scientific community".[283]

[279] ALBUQUERQUE *supra note* 148 at 879.
[280] SOUSA, Sandra Oliveira – *"It's all in your head?" – A Utilização Probatória de Métodos Neurocientíficos no Processo Penal"* in XX Estudos Comemorativos dos 20 Anos da FDUP. Vol II. 2017, p. 746.
[281] Daubert vs Merrel Dow Pharmaceuticals, Inc. 1993. *See* SOUSA, Susana Oliveira – *"It's all in your head?" – A Utilização Probatória de Métodos Neurocientíficos no Processo Penal"* in XX Estudos Comemorativos dos 20 Anos da FDUP. Vol II. 2017, p. 746; SOUSA, Susana Aires - *"Neurociências e Processo Penal: Verdade ex machina?"* in Estudos em Homenagem ao Prof. Doutor Manuel da Costa Andrade, vol. II, 2017, p. 895, 896.
[282] SOUSA, João Henrique Gomes de – *"A Perícia Técnica ou Científica Revisitada numa Visão Prático-Judicial"* in Julgar, n.º 15, 2011, p. 43 – 45.
[283] NUTTER, Patrick – "*Machine Learning Evidence: Admissibility and Weight"* in Journal of Constitutional Law, vol.21:3, 2019, p. 932.

In the case of AI generated evidence the ML algorithm and models in use must be supported by a valid scientific theory.[284] From the perspective of the admissibility of ML generated evidence the presented solution is feasible to mitigate the risk of automation bias, implying an active analysis and ponderation by the judges and participants regarding the reliability of the evidence.

Even though advanced drowsiness and attention detection systems are AI consumer products, therefore the studies and testing of each specific model demonstrating the machine's error rate will not probably be published nor peer reviewed, due to trade secrecy.[285] The underlying ML algorithms of drowsiness and attention detection systems (ANNs, FIS and SVM) consist of well-known techniques, enjoying a general acceptance in the field of scientific community, and peer-reviewed literature and research regarding such specific techniques has proliferated during the last years.[286] Besides, these systems have already been in use by automotive companies for years and some studies show that the accuracy level of these systems is quite elevated, in addition, sharing the same argument as Sabine Gless, these systems are a trusted mechanisms, otherwise, they would not be part of the mandatory advanced safety equipment in vehicles".[287]

Besides the ML algorithms behind its operation consist of already known scientific techniques which have been studied and tested in the last years, and drowsiness and attention detections are already used in some vehicles by automotive companies, consisting in recognized scientific methods.

## 4 The classification Problem

The general admissibility of an AI generated evidence will depend on the degree of complexity inherent to the algorithm which will vary according to the different levels of machine autonomy, its functioning, and goals. This means different AI systems will generate different levels of opacity and affect fundamental rights in different degrees, and therefore they require and *in casu* assessment.

[284] RODRIGUES supra note 229 at 22 and ALBERGARIA supra note 207 at 34-35.
[285] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 245
[286] See generally: BERGASA, Luis M. et al – *"Real-Time System for Monitoring Driver Vigilance"* in IEEE Transactions on Intelligent Transportation Systems, vol. 7, n. º 1, 2006.
[287] NUTTER supra note 297 at 933 and GLESS supra note 299 at 224.

After considering the first obstacles regarding the compliance with fundamental rights, it is time to understand if AI-generated machine evidence would be suitable to integrate any of the typical evidence frameworks in the Portuguese Code of Criminal Procedure, and if any of them is suitable to provide the required transparency to allow the respect for the rules of a fair trial.

### 4.1 Testimony

Testimonial evidence is regulated in art. 128º PCCP. The first rule refers to its object and limits, stating that a witness may only be heard on the facts (s)he directly knows and that constitute the object of evidence. According to art. 131º, n.º 1 PCCP any person not lacking legal capacity due to mental disorder can be a witness and may only refuse to testify in cases provided by law.

The testimony evidence and the witness role have been a cornerstone of the evidentiary system since Roman law and is considered "the eyes and ears of justice"[288], a privileged source of information to form a conviction in the judge about the accuracy of facts.[289]

However, what happens when the witness is not a human being anymore?[290]

In this context, Andrea Roth coined the term *machine testimony*, assuming that some machines can do what a human witness does, this is, make claims that serve as a source of truth to factfinders.[291] Hence, we must distinguish mere tools that assist humans in conveying information (e.g., traditional mechanical reproductions and electronic evidence in art. 167º PCCP) from intelligent machines capable to convey their own messages in a way it may implicate varying levels of credibility testing to assert its probative value.[292] In the case under analysis an AI system embodied in a vehicle will convey the message if the driver at the time of the accident showed signs of fatigue or distraction, assuming the equivalent role of an eyewitness.[293] Nevertheless, machine

[288] BENTHAM Apud TRIUNFANTE, Luís Lemos, et al – "*Comentário Judiciário do Código de Processo Penal*". t.II. (art. 128º). 2019, p. 86
[289] TRIUNFANTE, Luís Lemos, et al. – "*Comentário Judiciário do Código de Processo Penal*". t.II. (art. 128º). 2019, p. 86.
[290] CHENG, Edward – *"How to cross-examine a machine in court".* [Online].2016.[Last access: 17.05.2022]. Available at: https://news.vanderbilt.edu/2019/03/27/how-to-cross-examine-a-machine-in-court/
[291] ROTH, Andrea – "*Machine Testimony*" in Yale Law Journal, Vol. 126, n.º 1, 2017, p. 2040 – 2051.
[292] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 212.
[293] *Id* at 214.

testimony will not fit the traditional legal molds of testimony evidence as the way it is regulated was exclusively thought for human witnesses.

The human nature of testimonial evidence results from its own legal framework, being clear in some specific provisions, such as art. 131º PCCP regarding the capacity and duty to testify, and art. 132º PCCP which includes a list of duties and rights that may only be attended by humans, such as the right to take an oath when heard by judiciary authorities and answer with the truth. Besides, narrow AI does not allow sufficient scrutiny of the source of its assertions, meaning that they cannot undergo essential acts such as confrontation in art. 146º PCCP, witness examination and cross-examination under art. 138º PCCP and 348º PCCP. These systems are not capable to express themselves and explain the reasons behind their reasoning. For instance, if the witness is a human passenger that was in the car, (s)he could testify and be confronted about a defendant's driving ability, potential biases, misjudgment, or even intentional lying,[294] none of this is possible when the witness is the drowsiness detection system.

Thinking that AI-driven devices should undergo a similar credibility testing as witnesses because of their design and operationality, improperly places such machines on a similar footing as human witnesses,[295] as firstly it disregards the own specificities of the regulated testimonial evidence, amongst the above mentioned, the principle of immediacy, which, besides working as a safeguard for both defence and prosecution/assistant, serves as a tool for the judge to evaluate the facts properly. The principle of immediacy is enshrined in art. 355º PCCP, it implies the general rule that the court may only evaluate evidence that was produced and examined during the trial hearing. According to Dá Mesquita, this principle is focused on a relational dimension between the trier of fact with the evidence, meaning the judge must have the most immediate contact with the evidence.[296]

Second, machine evidence is not flawless and may suffer from infirmities the same way human witnesses do, such as biases, and analytical errors derived from wrong human

[294] Id at 222

[295] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials"*. 2020, p. 222

[296] DÁ MESQUITA, Paulo – *"Alguns Sinais Sobre Tendências Actuais do Processo Penal Português – Divergências Metodológicas sobre o Contraditório, a Prova, a Imediação e a Confiança nos Juízes"* in Julgar nº 25. 2015, p. 124. And SILVA, Germano Marques – *"Curso de Processo Penal",* Tomo I. 2010, p. 105.

operation, inputs, or malfunctioning.[297] The machine testimony requires a level of scrutiny and transparency, "especially with respect to the design, algorithms and machine learning/training data"[298] which cannot be achieved with the regime of traditional testimony evidence and would result in a severe hampering of the defendant's right to challenge, impeach, and confront the presented evidence, harming the right to contradictory and a fair trial, 32º/5º *in fine* PCR and art. 6º ECHR. In order for the defendant to challenge the accuracy of the machine statement they will need to understand how the machine was programmed and understand its reasoning models, to assess what kind of conditioning and how they are weighted by the machine to render that result.[299]

### 4.2 Hearsay Rule

The opacity problem inherent to this type of evidence has been generally compared to the concerns raised regarding anonymous witness.[300] According to jurisprudence of the European Court of Human Rights an anonymous witness is not necessarily incompatible with the right to a fair trial if the defendant can counterbalance the burden of anonymity.[301] The ECHR set three requisites that should be satisfied: there has to be a good reason for admitting the witness' absence (death or attributable fear); the conviction should not be solely based in the absent witness testimony; there has to be sufficient counterbalance to admit a fair and proper assessment of the reliability of the evidence to take place.[302]

It has been a practice to apply to AI systems the hearsay rule, in fact "courts when confronted with opaque machine evidence shoehorn them into existing rules by treating them as hearsay".[303] In the Portuguese criminal procedure, the most similar rule to hearsay

[297] ROTH, Andrea – "ROTH, Andrea – "*Machine Testimony"*. 2017. Yale Law Journal, Vol. 126. Nº.1,p.1993-2000.
[298] *Id.*
[299] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 231
[300] FIDALGO, Sónia – *"A Utilização de Inteligência Artificial no Âmbito da Prova Digital – direitos fundamentais (ainda mais) em perigo"* in A Inteligência Artificial no Direito Penal. 143 and ZAVRŠNIK, Aleš – *"Criminal Justice, Artificial Intelligence Systems, and Human Rights"* in ERA Forum 20, 2020, p. 576
[301] Refer to Al-Khawaja & Tahery v. U.K., Eur. Ct. H.R., App. No. 26766/05 & 22228/06, 37 (2011).
[302] Refer to Al-Khawaja & Tahery v. U.K., Eur. Ct. H.R., App. No. 26766/05 & 22228/06, 37 (2011) and GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 232, 233.
[303] ROTH, Andrea – "ROTH, Andrea – "*Machine Testimony"*,2017, Yale Law Journal, Vol. 126, n.º1 ,p. 1972

testimony is the indirect testimony *(depoimento indireto)* found in the provision of art. 129. PCCP.[304] The general rule is that this type of evidence is not admissible, whenever a witness' testimony is the result of what they heard another person say, the judge "may, *rectius*, must call such person to testify" [305] otherwise that part of the testimony cannot be used as evidence (art. 129º/1, 1st part PCCP). In the same way, there are hearsay dangers for humans, when they are not subject to oath, confrontation, and cross-examination, the same derives from the black box inherent to machine evidence.

In this case, the machine conveyances can be considered as hearsay assertions from the programmer. Therefore, to accept machine evidence the programmer should be called to testify.[306] Such solution does not seem feasible, at least for two reasons. The process of creating an AI program is complex, algorithms' design and establishing models of training data require the participation of different computer engineers and programmers, generating a context of distributed cognition in which none may explain for sure what lead to the resulting output. Also, the own functioning of machine learning technology envisages providing autonomy to the machine, this is, providing the machine with the capability to collect data, learn and perfect its decision according to its interactions with the surrounding environment, which means at some point none of the programmers might be able to explain why a decision was taken. Adding to these reasons, Gless points out that when vetting human testimony, the defendant wants to know what factors were perceived and to what measure they were considered and how they led to a particular conclusion,[307] the same is expected when vetting the machine.

According to Roth, there is an additional risk in calling the programmers as in these cases what is in cause is the functioning of a consumer product to which designers and programmers were paid to develop, resulting in a biased testimony.[308]

The indirect testimony is only accepted in specific cases predicted by law: when it is impossible to call the person-source of the original information to testify due to death,

[304] See generally: PINTO, Frederico de Lacerda Da Costa - *«Depoimento Indirecto, Legalidade da Prova e Direito de Defesa»,* in Estudos em Homenagem ao Prof. Doutor Jorge de Figueiredo Dias, 3.º Vol., Coimbra Editora, 2010.
[305] ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem."* 2018, p. 360
[306] ROTH, Andrea – "*Machine Testimony"* in Yale Law Journal, vol. 126. 2017, n.º 1, p. 1986.
[307] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials".* 2020, p. 233
[308] ROTH, Andrea – "*Machine Testimony"* in Yale Law Journal, vol. 126, 2017, n.º 1,p. 2036.

supervenient psychic anomaly and impossibility to be found (art. 129º/1º *in fine*). This solution is an exception to the principle of immediacy and favours the inquisitorial aspect of the criminal proceeding and the assessment of the material truth. The fact is that even without the testimony-source of the information, the indirect testimony may still be confronted and challenged by the defendant and produce a degree of conviction in the judge. This solution for its exceptional character does not allow analogy to other situations than the ones provided for in the law.[309]

### 4.3 Expert Evidence

Another perhaps more acceptable solution is to treat this evidence as requiring an expert intervention. Whenever the understanding and perception of fact requires technical, scientific, or artistic knowledge an expert must be nominated by judiciary authorities, art. 151º PCCP. This means the Portuguese criminal proceeding similarly to other European jurisdictions, such as France and Germany, does not admit what is called "expert-witness" that should be chosen and paid by the parties.[310]

At the first level of credibility testing, explaining the general techniques adopted by the system and machine's physical conditions experts may play an important role, and from this point of view will always be necessary to assess the validity of scientific techniques in use, and explain them to the trier of fact to achieve sufficient understanding and trustworthy fact-finding.[311]

One positive aspect from expert evidence is that they provide the impartiality that the machine's programmers could not have, nevertheless at a second stage of analysis will be affected on the same way by the effects of the black box. The expert's opinion will be the result of a distributed cognition between its own expertise, the programmers and the machine's expertise as well.[312]

---

[309] ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem",* 2018, p. 362

[310] SILVA DIAS, Maria do Carmo, et al. – "Comentário Judiciário do Código de Processo Penal". t.II. (art. 128º). 2019, p. 368.

[311] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials".* 2020, p. 212: "Machine evidence and expert testimony are inextricably linked due to the fact that AI-generated data must be explained. "

[312] ROTH, Andrea – "*Machine Testimony"* in Yale Law Journal, Vol. 126. 2017. Nº.1,p. 1978

The main question is whether the explanations and validation provided by the experts will translate into a sufficient level of transparency that allows the defendant to effectively challenge the machine generated evidence. According to Quattrocolo, it is necessary to establish a difference between explanation and justification and to comply with the requirement of a sufficient explanation: a clear mathematic language that allows an *ex post* reviewer to understand how the process evolved from the input to the output.[313] However, in order to achieve such knowledge of how the process evolved to a certain output it would require the expert to access the code of the machine and its specific reasoning models,[314] which will probably be under trade secrecy and may vary from automobile brand and vehicle models. Hence, according to Gless, experts encounter limitations in comprehensibly explaining how an AI-driven device evaluates the human driver or demonstrates a clear chain of causality.[315]

Even if experts had access to the code, AI expertise and ML complexity exceed human capacity. It would be difficult even when given adequate time to question and evaluate AI to accomplish a full explanation of all details of the operational process and conclusions.[316] This may lead to the conclusion that a residual level of opacity might be inevitable when trying to explain AI functioning, that even an expert will not be able to surpass.

### 4.4 The need for amending the Portuguese evidentiary system

AI generated machine evidence is part of a new generation of evidence for which the traditional human-centered evidentiary framework might not be ready to regulate. Therefore, new mechanisms are necessary to contextualize and satisfactorily explain the results of machine evidence, in order to admit the proper exercise of the defendant's rights. Following Roth and Gless conclusions, AI has a unique status that must be acknowledged in order for its message to be accurately visible to the parties, court and public.[317]

---

[313] QUATTROCOLO, Serena – "*An Introduction to AI and Criminal Justice"* in Revista Brasileira de Direito Processual Penal. Vol.5. nº3. 2019, p.1529

[314] ROTH, Andrea – "*Machine Testimony"* in Yale Law Journal, Vol. 126. 2017. Nº.1,p.2034:"Some experts have argued that access to the source code is the only meaningful way to determine whether a complex algorithm's method is both reliable and reliably applied."

[315] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 211

[316] *Id* at 240.

[317] *Id* 248. *And* ROTH, Andrea – "*Machine Testimony"* in Yale Law Journal, vol. 126. 2017, n.º 1, p. 2053.

One essential aspect is that machine learning will always comprise a level of opacity, therefore admitting machine testimony as a new type of evidence must imply a reinforcement of the right to defence. Admitting the level of opacity inherent to AI, would not be something new to Portuguese criminal procedure as it already admits under some circumstances the indirect testimony (art. 129º PCCP), which is revested of a certain degree of opacity, perhaps it could open another exception for machine evidence. Otherwise, assuming that the residual opacity of machine evidence is not accepted under any circumstances will ban AI generated evidence for good from Portuguese criminal procedure and close space for future debate, also harming the principle of investigation by disowning an effective mean of truth assessment.

The indirect testimony accepts a level of inscrutability of the information source,[318] that is mitigated through the analysis of the indirect testimony. Likewise, AI systems opacity could be accepted by assuring ways of providing a sufficient degree of disclosure regarding its functioning. In this regard, Završnik understands that the same logic applied to the anonymous witness, should be applied to AI systems, under the condition that a fair balance between the right to participate effectively in trial and the use of AI systems in favour of truth assessment would be granted.[319]

As aforementioned it seems the appropriate method to provide some degree of disclosure regarding AI systems functioning would be the expert evidence, as according to the Portuguese criminal procedure experts are the suitable method of clarifying and translating scientific knowledge to the court. Nevertheless, this will require a reinforcement of the right to defence, already approached by David Ramalho with regards to digital evidence admissibility and Susana Sousa e Oliveira about scientific evidence.[320] Besides the right to challenge the output generated by the machine, the defendant must also have the right to effectively challenge the credibility of the evidentiary method itself, and the experts' report.

[318] ROTH, Andrea – "*Machine Testimony"* in Yale Law Journal, vol. 126. 2017, n.º 1, p. 1978.
[319] ZAVRŠNIK, Aleš – *"Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings",* 2019, p. 576, 577.
[320] RAMALHO, David Silva – *"Métodos Ocultos de Investigação Criminal em Ambiente Digital"*, 2017, p. 260, 261 and SOUSA, Sandra Oliveira – *"It's all in your head? – A Utilização Probatória de Métodos Neurocientíficos no Processo Penal"* in XX Estudos Comemorativos dos 20 Anos da FDUP. Vol II. 2017, p. 748.

Expert evidence is the only evidence in the portuguese criminal code procedure that is subtracted from the judge's free valuation, in accordance with the presumption from art. 163º, n. º 1 PCCP, also if the judge disagrees with the expert's report it must substantiate the reasons behind his disagreement, comprising an additional burden to discredit the expert opinion (n.º 2). This is a result from the traditional conception that science is an absolute truth, capable of offering unlimited, unfailing responses.[321]As aforementioned, AI systems proof the contrary, that is not even experts will be able to unveil AI's black box. The experts' report shouldn't result in an automatic acceptance of evidence reliability, let alone when referring to a type of evidence that comprises residual opacity and inscrutability. If it is the gate to the entrance of AI systems in the proceeding it must be given the defendant's an effective right to challenge the expert's report and opinion. It is also important to avoid the risk of machine evidence to be treated as a pure expert evidence, when it is in fact a machine testimony translated by experts.[322]

Taking the above into consideration it is here proposed some requirements to treat AI generated evidence:

1) The AI system under analysis must admit sufficient scrutiny. It is known that AI expertise exceeds human capacity, nevertheless it should be possible for the expert to elaborate an exhaustive report[323] that allows the participants in the criminal procedure and judge "to clearly understand how the machine gathers information, evaluates it, and how it makes a determination".[324] Reference is made to a "functional evaluation, that does not necessarily involve the same explanations as in human testimony".[325]

This means that due to its functioning and opacity level some AI systems will not be considered scrutable enough to be admitted as evidence, at least at the current stage of scientific evolution.

[321] SOUSA, Sandra Oliveira – *"It's all in your head? – A Utilização Probatória de Métodos Neurocientíficos no Processo Penal"* in XX Estudos Comemorativos dos 20 Anos da FDUP. Vol II. 2017, p. 749.
[322] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 237
[323] About the requirements to an exhaustive report refer to RAMALHO, David Silva – *"Métodos Ocultos de Investigação Criminal em Ambiente Digital"*, 2017, p. 258.
[324] GLESS, Sabine – *"AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials",* 2020, p. 240.
[325] *Id.*

2) Similarly, to was it has been decided by ECHR[326] regarding anonymous witnesses it must be assured that the AI generated evidence is not the sole or decisive source to generate the court's conviction.

3) Likewise, a fair counterbalance between the right of defence, and the truth assessment provided by the AI system must be granted, so it is possible to sufficiently cross-examine the algorithm functioning, the considered data and accuracy rate of the decision.

This would imply changes in the current model of expert evidence, to allow an effective impeachment of the produced evidence and expert opinion. A full legal reform to the already set legal framework would cause "devastating effects" and a profound legal reform that "would affect the efficiency of truth assessment".[327] It also important to avoid converting the experts in expert-witnesses, as they are criticized for the risks of resulting in expert partiality, negative economic impact for the participants resulting in an accentuated unbalance based on their economic power, *expert shopping* and lack of scientific reliability.[328]

As already suggested among Portuguese scholarship this would pass by providing a more significant role to technical consultants *(consultores técnicos)*, art. 155º PCCP, besides their limited participation of assisting experts and providing non biding opinion.

## 5. The European Commission's solution: Human Rights by Design

It is a fact that AI brings efficiency and is providing solutions to different challenges of modern society, nevertheless, it has a twofold effect. As while bringing benefits to our lives and to specific sectors, namely criminal justice, they simultaneously raise a set of technical, ethical, and juridic concerns.[329]

With regards to juridic concerns, AI systems's characteriscs may endanger fundamental rights when applied at the service of criminal justice, mainly regarding the defendant's

[326] Refer to Al-Khawaja & Tahery v. U.K., Eur. Ct. H.R., App. No. 26766/05 & 22228/06, 37 (2011).
[327] SOUSA, João Henrique Gomes de – *"A Perícia Técnica ou Científica Revisitada numa Visão Prático-Judicial"* in Julgar, n.º 15, 2011, p. 31.
[328] *Id* at 30.
[329] See SMUHA, Nathalie A. – *"The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence"*. 2019, p. 1. Available at SSRN: https://ssrn.com/abstract=3443537

fundamental rights (about this subject refer to section 4.1.). The solution of the European Commission in October 2017 the European Council recognized the sense of urgency to address emerging trends, in special AI, while ensuring a "high level of data protection, digital rights and ethical standards",[330] by inviting the Commission to "put forward a European approach to artificial intelligence".[331]

In order to achieve such goal, the Commission supports a human-centric approach of AI that takes into account the Ethic Guidelines prepared by the HLEG, while promoting the creation of a regulatory AI framework.[332] A human-centric approach to AI places people "at the centre of AI development",[333] considering AI systems as tools at the service of human well-being. This means human rights, including the ones presented in the Treaties of European Union, EU Charter of Fundamental Rights, and ethical guidelines as well, must be the central pillars in AI systems from their design and learning phases, to their deployment and several uses.[334]

In this context, the EU published in 2018 its communication for AI where it promoted the development of AI technology whilst emphasizing the need to ensure an appropriate ethical and legal framework, based in the EU's founding values and the EU Charter of Fundamental Rights.[335] In 2019 the High Level Expert Group on AI published the ethical guidelines to create a trustworthy AI.[336] According to the HLEG there are three essential components to achieve a trustworthy AI: 1) AI systems must be lawful, meaning they must comply with law and applicable regulations; 2) must be ethical, following ethical principles and values; 3) at last, an AI system must be technically and socially robust, meaning such systems must perform in safe, reliable manners so that society may trust that AI will not cause unintentional harm.[337]

---

[330] EUCO 14/17. Brussels, 19.10.2017. – *"European Council Meeting",* p. 7.
[331] *Id.*
[332] *Id.* at 9.
[333] COM(2019) 168 final. Brussels, 08.04.2019 - *Building Trust in Human-Centric Artificial Intelligence,* p.1.
[334] European Commission for the Efficiency of Justice (CEPEJ) – *"European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment"*.2018, p. 8 *and* COM(2019) 168 final. Brussels, 08.04.2019 - *Building Trust in Human-Centric Artificial Intelligence,* p.2.
[335] COM (2018) 237 final. Brussels, 25.04.2018 – "*Artificial Intelligence for Europe"*, p.13
[336] *See generally*: FLORIDI, Luciano – *"Establishing the Rules for Building Trustworthy AI"*. 2019. Available at SSRN: https://ssrn.com/abstract=3858392
[337] High Level Expert Group on Artificial Intelligence – "*Ethical Guidelines for Trustworthy AI."* 2019, pp.5-7.

The HLEG approach focuses on the last two components of ethical and robust AI, presenting four ethical principles rooted in the respected for respect for human dignity, mental and physical integrity: *Respect for human autonomy*, when interacting with AI, humans must keep full effective self-determination, and have meaningful opportunity for choice; *Prevention of harm*, AI systems should neither cause, exacerbate harm or adversely affect human beings. Also, they must be technically robust, and the environment in which they operate must be safe and secure. *Fairness*, AI systems must ensure an equal and just distribution of benefits and costs, respecting the principle of proportionality when balancing competing interests and objective. They must ensure as well that individuals and groups do not suffer unfair bias, discrimination, and stigmatization, and allow the ability to effectively contest its decision-making processes; *Explicability*, the decision-making processes of AI systems must be transparent to the possible extent to those directly and indirectly affected.[338]

Besides the outlined principles, the HLEG proposes seven key requirements that AI stakeholders (developers, deployers and end-users) must fulfil to achieve a trustworthy AI: *human agency and oversight; technical robustness and safety; Privacy and data governance; Transparency; Diversity, non-discrimination and fairness; Societal and environmental wellbeing; Accountability.*[339]

The HLEG was not the first to establish an ethical framework for AI.[340] In this regard, it takes special interest the European Commission for the Efficiency of Justice (CEPEJ) Ethical Charter on the use of AI in judicial systems and their environment,[341] which adopted 5 fundamental principles that must lead the deployment of AI systems at the service of justice systems: [342]

[338] I*d* at 12, 13.
[339] I*d* at 14.
[340] SMUHA *supra note* 296 at 4.
[341] European Commission for the Efficiency of Justice (CEPEJ) – *"European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment"*, 2018.
[342] *See generally*: ROSÁRIO, Rita Alexandre - *"Ética e Inteligência Artificial no Conselho da Europa"* in Anatomy of Crime. Journal of Law and Crime Sciences, n.º12, 2020, pp. 134 - 167.

*Principle of respect for fundamental rights*, this means "a preference should be given to *human-rights-by-design approaches"*.[343] The regulation regarding the conformity of AI design, programming, deployment and use, must prohibit direct or indirect violations of fundamental rights. Also, judicial decisions based on AI tools must fully comply the fundamental rights guaranteed in the ECHR, mainly the right to a fair trial, and the Convention on the Protection of Personal Data.

*Principle of non-discrimination*, some AI systems algorithms may comprise a risk of bias due to data grouping and classification relating to individuals and specific traits referring to groups of individuals, particular care should be taken at the designing and deployment phases, to effectively detect possible biases and adopt suitable corrective measures; *principle of quality and security*, the machine learning models and collected data must derive from certified and qualified sources, this also refers to system integrity and intangibility evolving the need for storage and execution on safe environments;

*Transparency, impartiality and fairness* this principle might be a lead to solve future problems regarding AI generated evidence opacity. It demands a balance between intellectual property of certain processing methods and the need to access to the design process (transparency), in order to assess any possible bias (impartiality), and fairness and intellectual integrity. In these cases, the interests of justice must be prioritized, as AI systems may take a significant role in affecting people's lives.

*Principle "under user control"* the use of AI systems should not restrict human autonomy, it must by the contrary increase it. With regards to AI tools assisting decision-making process, they should not necessarily bound the professionals in the judicial systems to a decision. On the particular user's perspective they should clearly inform if the solution is binding and inform that the user has the right to legal advice and the right access to a court.

Even though the proliferation of ethical principles is important, as to their persuasive nature they may influence and stimulate debate within decision making, they consist of

[343] European Commission for the Efficiency of Justice (CEPEJ) – *"European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment"*, 2018, p. 8.

soft law guidelines.[344] Ethical guidelines alone aren't sufficient to construct an environment of lawful AI, mainly regarding the need to establish effective rules when developing and deploying AI systems in the context of a single European market and criminal justice systems.

The European Commission is already engaged in creating a lawful environment with the attempt of presenting a horizontal regulation on AI with the presentation of an AI Act Proposal. Nevertheless, some criticism has been made for the proposed solutions, mainly referring to the suggested AI definition, and its application scope regarding the allowed exceptions and gaps (mainly regarding social scoring and biometric identification systems), the lack of robustness of *ex-ante* and *ex-post* requirements for AI conformity assessment, and lack of responsibility allocation for the purposes of fundamental rights protection. LEADS members point out that the current draft of the Proposal "does not ensure and effective framework for the enforcement of legal rights and responsibilities, nor does it provide sufficient protection to maintain the rule of lar and democracy".[345]

There still is a long path for European Union to cross in what regards to AI systems control, regulation, and integration in juridic systems.

## 6. Conclusions

During the last years AI driven technology became ubiquitous in our lives, pervading high-impact sectors and changing the way we live. The current paradigm of AI consists of software systems designed by humans, that act in the digital or in the physical world (through hardware) with learning rationality, to achieve a specific goal. To achieve rationality, these systems must be able to perceive their environment, collect and interpret data, adapt and reason on the knowledge obtained from the collected data deciding the most suitable action to take. Such capabilities are enabled by reasoning/decision-making techniques and machine learning. The current stage of scientific evolution is "limited" to narrow AI, non-self-conscient systems designed to perform specific tasks autonomously, without or with little human intervention.

[344] CAHAI (2020) 07-fin. Ad Hoc Committee on Artificial Intelligence (CAHAI) – "AI Ethics Guidelines: European and Global Perspectives". 2020, p.5.

[345] SMUHA, Nathalie, et al – *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal For an Artificial Intelligence Act*. LEADS Lab University of Birmingham. 2021, p. 9. Also refer to the criticism presented by EBERS, Martin, et al - *The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)*. Multidisciplinary Scientific Journal.

The dissemination of AI systems and given their capacity to monitor and autonomously evaluate human behaviour might generate potential new means of evidence, such as the drowsiness and alert detection systems that assess the drivers' capacity to drive. Its inherent opacity and inscrutability generate the black box paradox which makes them potentially harmful for structuring principles of criminal proceedings and fundamental rights, mainly the right to a fair trial in the dimension of the plain exercise of the right to defence.

AI generated evidence challenges the traditional human-centric evidentiary framework and requires new solutions for its admissibility that goes beyond the traditional human testimony, hearsay and the classic method of expert evidence. Regarding the Portuguese criminal procedure, it is safe to say that the current framework isn't ready to receive AI system is ways to ensure an effective right of defence and to a fair trial.

Even though the Portuguese criminal procedure is open to admit new means of evidence besides the ones already typified by law (atypical evidence), an effect from art. 125º CPP, it is necessary to assess the admissibility of the atypical evidence in light of the legal limit of the respect for fundamental rights, subsidiarity regarding the already regulated evidence in order to avoid subversions of their existing rules, and their relevance and reliability.

A conclusion to take is that AI systems admissibility in courtrooms will vary, depending on their autonomy level, opacity and functioning. This means AI systems that comprise unacceptable risk to fundamental rights, nor allow a sufficient scrutiny of its decision-making processes, not providing a sufficient explanation for the defendant, court and participants will not be admissible. This also means that some AI systems might be in different stages of the evidentiary lifecycle, some of them being *to new to be reliable,* while others might be eligible to be *subject to testing and regarded as generally reliable.*

After an analysis of the pre-existing evidentiary rules from Portuguese criminal procedure another conclusion to take is that AI generated assertions as the ones resulting from advanced drowsiness and attention detection system for its specificities, resulting a machine testimony that requires innovative measures to be admitted in courts.

A sign that their admissibility would not be an impossibility is that the Portuguese criminal procedure already accepts a level of opacity under the conditioned admissibility of the indirect testimony. In this regard, ECHR jurisprudence has been flexible in considering that opacity resulting from the admissibility of the testimony of an anonymous witness isn't necessary incompatible with the defence rights, if sufficient counterbalance between truth assessment and the possibility to challenge the evidence is provided.

The expert evidence might be a gate of entrance of AI generated evidence in the Portuguese criminal procedure to validate the scientific method applied and explain the decision-making process. In such case it will be necessary to reinforce the right to defence, to allow a proper challenge of the expert's report and opinion, the credibility of the evidentiary method itself and the generated output. The suggestion solution would pass by attributing to technical consultants a significant role to effectively impeach the expert report.

# Bibliography:

## Authors:

ALBERGARIA, Pedro Soares, et al. – "*Comentário Judiciário do Código de Processo Penal"*, tomo. II, (art. 125º), Almedina, 2019.

ALBUQUERQUE, Paulo Pinto – "*Comentário do Código de Processo Penal à luz da Constituição da República e da Convenção Europeia dos Direitos do Homem",* 4ª edição, Lisboa, Universidade Católica Editora, 2018.

An Executive's guide to AI: *Why AI now?* Available at: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai

An Executive's guide to AI: *Why AI now?* Available at: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai

BARTRAM, Robert, et al – "*The Age of Artificial Intelligence"*. ISO Focus Magazine. Nov-Dec 2019. Available at: https://www.researchgate.net/publication/350622618_Discrimination_in_the_age_of_artificial_intelligence

BEER, Jenay M; FISK, Arthur. D; ROGERS, Wendy. A – "*Toward a framework for levels of robot autonomy in human-robot interaction"* in Journal of Human-Robot Interaction, n. º 3(2):74, 2014.

BEN-ISRAEL, Issac, et al. – "*Towards Regulation of AI Systems- Global Perspectives on the development of a legal framework on Artificial Intelligence systems based on the Council of Europe's standards on human rights, democracy and the rule of law"*, 2020.

BERGASA, Luis M., et al. – *"Real-Time System for Monitoring Driver Vigilance"* in IEEE Transactions on Intelligent Transportation Systems, vol. 7, n. º 1.

BOMHARD, David; MERKLE, Marieke – *"Regulation of Artificial Intelligence. The EU Commission's Proposal of an AI Act"* in EuCML, Issue 6, 2021, p. 260.

BORGESIUS, Frederik Zuiderveen – *"Discrimination, Artificial Intelligence, and Algorithmic Decision Making",* Council of Europe, 2018.

BOSTROM, Nick – "*Superintelligence: Paths, Dangers, Strategies"*, Oxford University Press, 2014.

BOUCHER, Philip – *"Artificial Intelligence: How does it work, why does it matter, and what can we do about it".* Study for the Panel for the Future of Science and Technology, 2020.

BOUCHER, Philip – *"What If We Chose New Metaphors For Artificial Intelligence?".* European Parliamentary Research Service, 2021.

BURREL, Jenna – *"How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms"* in Big Data and Society (I-12), 2016.

CARIA, Rui – *"O Caso State v. Loomis – A Pessoa e Máquina na Decisão Judicial"* in A Inteligência Artificial no Direito Penal, Coimbra, Almedina, 2020.

Cavaleiro de Ferreira, Manuel – "*Curso de Processo Penal"*, vol. I, Lisboa, Editora Danúbio, 1986.

CHESTERMAN, Simon – *"Trough a Glass, Darkly: Artificial Intelligence and The Problem of Opacity"* in NUS Law Working Paper nº 2020/1, 2020.

DÁ MESQUITA, Paulo – *"Alguns Sinais Sobre Tendências Actuais do Processo Penal Português – Divergências Metodológicas sobre o Contraditório, a Prova, a Imediação e a Confiança nos Juízes"* in Julgar, n.º 5, Coimbra Editora, 2015.

DÁ MESQUITA, Paulo, et al. - "*Comentário Judiciário do Código de Processo Penal"*, Tomo II, (art. 147º), Almedina, 2019.

D*IAS,* Jorge Figueiredo – *"Sobre os Sujeitos Processuais no Novo Código de Processo Penal"* in Jornadas de Direito Processual Penal: O novo código de processo penal, 1998.

DIAS, Jorge Figueiredo – "*Direito Processual Penal"*, vol. I, Coimbra, 1974.

DONG, Yanchao, et al – *"Driver Inattention Monitoring System for Intelligent Vehicles: A Review."* in IEEE Transactions On Intelligent Transportation Systems, vol. 12, n. º 2, 2011.

DUARTE, Eurico Balbino – *"Making Of – A Reconstituição do Facto no Processo Penal Português"* in Prova Criminal e Direito de Defesa: Estudos sobre a Teoria da Prova e Garantias da Defesa em Processo Penal, Coimbra, Almedina, 2013.

EBERS, Martin, et al – "*The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)"* in Multidisciplinary Scientific Journal, n.º 4, 2021.

ENZWEILER, Markus – "*The Mobile Revolution – Machine Intelligence for Autonomous Vehicles"* in IT – Information Technology, vol. 57, 2015.

FIDALGO, Sónia – *"A Utilização de Inteligência Artificial no Âmbito da Prova Digital – Direitos Fundamentais (ainda mais) em Perigo"* in Inteligência Artificial no Direito Penal, Coimbra, Almedina, 2020.

FLORIDI, Luciano – "*Distributed Morality in an Information Society"* in Science and Engineering Ethics, n.º 19, 2012.

FLORIDI, Luciano – "*Should we be afraid of AI?"* in AEON Magazine, 2016. Available at: https://aeon.co/essays/true-ai-is-both-logically-possible-and-utterly-implausible

FLOWERS, Johnathan – "*Strong and Weak AI: Deweyan Considerations" in* AAAI Spring Symposium: Towards Conscious AI Systems, 2019.

GIALUZ, Mitja – "Quando la Giustizia Penale Incontra L'Intelligenza Artificiale: Luci e Ombre Dei Risk Assessment Tools Tra Stati Uniti Ed Europa" in Diritto Penale Contemporaneo, 2019. Available at: https://archiviodpc.dirittopenaleuomo.org/d/6702-quando-la-giustizia-penale-incontra-l-intelligenza-artificiale-luci-e-ombre-dei-risk-assessment-too

GIUFFRIDA, Iria – *"Liability for AI Decision-Making: Some Legal and Ethical Considerations"* in William & Mary Law School Scholarship Repository, 2019.

GIUFFRIDA, Iria; LEDERER,Frederic; VERMERYS, Nicolas – "*A Legal Perspective on the Trials and Tribulations of AI: How Artificial Intelligence, the Internet of Things, Smart Contracts and Other Technologies Will Affect the Law"* in Case Western Reserve Law Review, vol. 68, 2018.

GLESS, Sabine – "*AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials"* in Georgetown Journal of International Law, vol. 51, n.º 2, 2020.

GLESS, Sabine, SILVERMAN, Emily, WEIGEND, Thomas – *"If Robots Cause Harm, Who is To Blame? Self-Driving Cars and Criminal Liability"*in New Criminal Law Review: An International and Interdisciplinary Journal, vol. 19, n.º 3, 2016.

GUGERTY, Leo – *"Newell and Simon's Logic Theorist: Historical Background and Impact on Cognitive Modeling"* in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2006. Available at https://www.researchgate.net/journal/Proceedings-of-the-Human-Factors-and-Ergonomics-Society-Annual-Meeting-1071-1813

HAENLEIN, Michael; KAPLAN Andreas, – "*Siri, Siri, in my hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence"* in Business Horizons, vol. 62, 2019.

HAENLEIN, Michael; KAPLAN, Andreas –*"A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence"*in California Management Review, n.º 61 (4), 2019.

HANSEN, John. H. L, et al. – "*Driver Modeling for Detection & Assessment of Distraction: Examples from the UTDrive testbed"*in IEEE Signal Processing Magazine, n.º 34 (4), 2017.

HERWEIJER, Celine – "*8 Ways AI can Help Save the Planet"* in World Economic Forum, 2018. Available at: https://www.weforum.org/agenda/2018/01/8-ways-ai-can-help-save-the-planet/

ISIDORE, Chris – "*Police say no one was in driver's seat in fatal Tesla crash* in CNN Business", 2021. Available at : https://edition.cnn.com/2021/04/19/business/tesla-fatal-crash-no-one-in-drivers-seat/index.html

JANUÁRIO, Túlio Xavier – "*Veículos Autónomos e Imputação de Responsabilidades Criminais por Acidentes* in A Inteligência Artificial no Direito Penal", Coimbra, Almedina, 2020.

JONES-CELLAN, Rory – *"Uber's self-driving operator charged over fatal crash",* 2020, BBC news available at: https://www.bbc.com/news/technology-54175359

KERR, Orin – "*Digital Evidence and The New Criminal Procedure"* in Colombia Law Review, vol. 105, 2005.

KORF, Richard E. – "*Does Deep-Blue use AI?"* in AAAI Technical Report, WS-97-04, 1997.

LIGETI, Katalin – *"Artificial Intelligence and Criminal Justice"* in AIDP-IAPL International Congress of Penal Law, 2019.

LIPTON, Zachary C. – *"From AI to ML to AI: On Swirling Nomenclature & Slurried Thought"*, 2018. Available at: https://www.approximatelycorrect.com/2018/06/05/ai-ml-ai-swirling-nomenclature-slurried-thought/

MANSON, Stephen; SENG, Daniel – *"Artificial Intelligence and Evidence"* in Singapure Academy of Law Journal, n.º 33, 2021.

MARQUES DA SILVA, Germano - *"Curso de Processo Penal II"*, 4ª edição, Editorial Verbo, 2008.

MCCARTHY, John – *"What is Artificial Intelligence"*, available at: whatisai.dvi (unimi.it)

MELCHER, Vivien, et al. – "*Take-Over Requests for Automated Driving."* In Procedia Manufacturing, n.º 3, 2015.

MESQUITA, Paulo Dá – "*A Prova do Crime e o que se disse Antes do Julgamento – Estudo sobre a Prova no Processo Penal Português à Luz do Sistema Norte-Americano*", Coimbra, Coimbra Editora, 2011.

MINSKY, Marvin – "*Semantic information Processing"* in MIT Press, 2015.

MOREIRA, Vital; CANOTILHO, Gomes – *"Constituição da República Anotada" ( artº 20)*, vol. 1, Coimbra Editora, 2007.

NEWELL, Allen and HERBERT, Simon – *"Computer Science as Empirical Inquiry: Symbols and Search",* 1976.

NILLS, Nilson J – *"Artificial Intelligence: A New Synthesis",* San Francisco, CA.: Morgan Kaufmann Publishers, 1998.

NILSSON, Nils J. – *"The Quest for Artificial Intelligence: A History of Ideas and Achievements"*, Cambridge University Press, 2010.

NUNN, Alexander – "*Machine Generated Evidence in The SciTech Lawyer: A publication of the American Bar Association" in* Science & Technology Law Section, vol 16, n.º 5, 2020, pp. 4, 5.

NUTTER, Patrick – "*Machine Learning Evidence: Admissibility and Weight"* in Journal of Constitutional Law, Vol. 21:3, 2019.

OLIVEIRA, Arlindo – "*Inteligência Artificial"*, Fundação Francisco Manuel dos Santos, 2019.

OLIVEIRA, Licínio – "Driver Drowsiness Detection Using Non-Intrusive Signal Acquisition", 2018. Available at: https://repositorio-aberto.up.pt/bitstream/10216/113802/2/276791.pdf

PAGALLO, Ugo – *"Research Handbook on the Law of Artificial Intelligence",* Edward Elgar Publishing, 2018.

PAGALLO, Ugo; QUATTROCOLLO, Serena – "*The Impact of AI on Criminal Law and its Twofolds Procedures"* in Research Handbook on the Law of Artificial Intelligence, Edward Elgar Publishing, 2018.

PARASURAMAN, Raja; SHERIDAN, Thomas. B; WICKENS, Christopher D. – "*A Model for Types and Levels of Human Interaction with Automation*" in IEEE Transactions on Systems, Man, and Cybernetics-Part A: Sistems and Humans", vol. 30, n.º 3, 2000.

PELERIGO, Vanessa – *"Brain Computer Interface – Uma Primeira Abordagem"* in Anatomy of Crime: Journal of Law and Crime Sciences, n.º 12, 2020.

POSZLER, Franziska; GEISSLINGER, Maximilian – "*AI and Autonoums Driving: Key Ethical Considerations"*, 2021, p. 2.

PRIEST, Colin – "*Humans and AI: Should we describe AI as autonomous?",* 2021. Available at Blog / AI & ML Expertise: https://www.datarobot.com/blog/humans-and-ai-should-we-describe-ai-as-autonomous/

QUATTROCOLO, Serena – "*An Introduction to AI and Criminal Justice"* in Revista Brasileira de Direito Processual Penal. vol.5, nº3, 2019.

RAMALHO, David Silva – *"Métodos Ocultos de Investigação Criminal em Ambiente Digital"*, Coimbra, Almedina, 2017.

RAPOSO, Vera Lúcia – "*Draft Regulation on Artificial Intelligence: The devil is in the details"* in Privacy and Data Protection Magazine, n. º3, 2021.

ROBALO, Inês –*"Verdade e Liberdade – A Atipicidade da Prova em Processo Penal",* 2012. Available at: https://repositorio.ucp.pt/bitstream/10400.14/15696/1/Verdade%20e%20Liberdade%20-%20A%20Atipicidade%20da%20Prova%20em%20Processo%20Penal%20-%20In%C3%AAs%20Robalo.pdf

RODRIGUES, Anabela Miranda – "*Inteligência Artificial no Direito Penal – A Justiça Preditiva entre a Americanização e Europeização* in A Inteligência Artificial no Direito Penal", Coimbra, Almedina, 2020.

ROTH, Andrea – "*Machine Testimony*", 2017, Yale Law Journal, Vol. 126, n.º 1

RUSSEL, Stuart; NORVIG, Peter – *Artificial Intelligence: A Modern Approach*, 2010, pp. 13, 14.

*SACHOULIDOU, Athina – "OK Google: is (she) guilty?"* in Journal of Contemporary European Studies, 2021.

SAE International's blog- "*SAE Levels of Driving Automation Refined for Clarity and International Audience"*, 2021

SAE J3016 Recommended Practice: *"Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles",* 2021.

SAMOILI, Sofia, et al. – "*AI Watch Defining Artificial Intelligence: Towards an operational definition and taxonomy of artificial intelligence"*. JRC Technical Reports. 2020.

SCHANK, Roger C. – "*What is AI, Anyway?"* in AI Magazine, vol. 8, n.º 4, 1987.~

SCHAWB, Klaus – The Fourth Industrial Revolution, EditPro, 2017.

SCHERER, U Mattew – *"Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies",* 2016, p. 360~

SEIÇA, ALBERTO MEDINA DE, "Legalidade da prova e Reconhecimentos «atípicos» em processo penal: notas à margem de jurisprudência (quase) constante", Liber Discipulorum para Jorge de Figueiredo Dias, organizado por Manuel da Costa Andrade, Coimbra, Coimbra Editora, 2003.

SHIN, Donghee – *" How do Users Interact with Algorithm Recommender Systems? The Interaction of Users, Algorithms, and Performance"* in Computer in Human Behaviour, vol. 109, 2020.

SILVA, Sandra Oliveira – *"Legalidade da Prova e Provas Proibidas" in* Revista Portuguesa de Ciência Criminal , n.º 4, 2011.

SMUHA, Nathalie A. – "The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence", 2019.

SMUHA, Nathalie, et al – "*How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal For an Artificial Intelligence Act".* LEADS Lab University of Birmingham, 2021.

SOUSA, João Henrique Gomes de – *"A Perícia Técnica ou Científica Revisitada numa Visão Prático-Judicial"* in Julgar, n.º 15, 2011.

SOUSA, Paulo Mendes – *"Lições de Direito Processual Penal",* Almedina, 2014.

SOUSA, Sandra Oliveira – *"It's all in your head?" – A Utilização Probatória de Métodos Neurocientíficos no Processo Penal"* in XX Estudos Comemorativos dos 20 Anos da FDUP. Vol II. 2017.

*SOUSA, Susana Aires* - *"Neurociências e Processo Penal: Verdade ex machina?" in Estudos em Homenagem ao Prof. Doutor Manuel da Costa Andrade,* vol. II, 2017.

SOUSA, Susana Aires de – *"Não fui eu, foi a máquina: Teoria do Crime, Responsabilidade e Inteligência Artificial"* in Inteligência Artificial no Direito Penal, 2020.

STONE, Peter, et.al – "*Artificial Intelligence and Life in 2030 - One Hundred Year Study on Artificial Intelligence". Report of the 2015 Study Panel*. 2016.

SURDEN, Harry; WILLIAMS, Mary-Anne – *"Technological Opacity, Predictability, and Self-Driving Cars"* in Cardozo Law Review. Vol.38. 2016.

The Royal Society Report - *"Machine Learning: The Power and Promise of Computers that Learn by Example"*, 2017.

TOTSCHNIG, Wolfhart – "*Fully autonomous AI"* in Science and Engineering Ethics, n.º 26(5), 2019

TRIUNFANTE, Luís Lemos, et al. – "*Comentário Judiciário do Código de Processo Penal"*, Tomo.II, (art. 128º), Almedina, 2019.

TURING, Alan. M – "*Computer Machinery and Intelligence*. Mind – A Quarterly Review of Psychology and Philosophy", 1950.

VALENTE, Manuel Monteiro Guedes – "*Processo Penal",* Tomo I, Almedina, 2020.

VEALE, Michael – *"Algorithms in the Criminal Justice System".* A Report by Law Society of England and Wales, 2019.

WANG, Pei – "*On the Working Definition of Intelligence"* in Center for Research on Concepts and Cognition Indiana University, 1995.

ZAVRŠNIK, Aleš – *"Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings".* 2019.

ZAVRŠNIK, Aleš – *"Criminal Justice, Artificial Intelligence Systems, and Human Rights"* in ERA Forum 20. 2020.

ZILBERSTEIN, Shlomo – "*Building Strong Semi-Autonomous Systems" in* Twenty-Ninth AAAI Conference on Artificial Intelligence, vol. 29, n.º 1, 2015.

**Documents:**

CAHAI (2020) 07-fin. Ad Hoc Committee on Artificial Intelligence (CAHAI) – "AI Ethics Guidelines: European and Global Perspectives". 2020.

COM (2015) 192 final. Brussels, 06.05.2015 – *A Digital Single Market Strategy for Europe*

COM (2018) 237 final. Brussels, 25.04.2018 – *Artificial Intelligence for Europe.*

COM (2020) 65 final. Brussels, 19.02.2020 – *White Paper on Artificial Intelligence – A European approach to excellence and trust.*

COM (2021) 206 final. Brussels, 21.04.2021 – Proposal for a Regulation of The European Parliament and of the Council. Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts.

COM(2018) 237, European Commission, 2018.

COM(2018) 293 final. Brussels. 2018, p.2. EUROPE ON THE MOVE Sustainable Mobility for Europe: safe, connected, and clean.

COM(2019) 168 final. Brussels, 08.04.2019 - *Building Trust in Human-Centric Artificial Intelligence.*

EUCO 14/17. Brussels, 19.10.2017. – European Council Meeting.

European Group on Ethics in Science and New Technologies – "*Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems",* 2018.

MSI-AUT(2018)05. Committee of Experts on Human Rights Dimensions of Automated Data Processing and different forms of Artificial Intelligence – "*A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework".*

Royal Academy of Engineering Report – "*Autonomous Systems: Social, Legal and Ethical Issues",* 2009.

The European Commission's High-Level Expert Group on Artificial Intelligence – "*A Definition of AI: Main Capabilities and Scientific Disciplines",* 2018.

The European Commission's High-Level Expert Group on Artificial Intelligence – *A Definition of AI: Main Capabilities and Scientific Disciplines*. 2018.

The European Commission's High-Level Expert Group on Artificial Intelligence – *A Definition of AI: Main Capabilities and Scientific Disciplines*. 2018.

**Jurisprudence:**

*AC. STJ, n.º 251/15.3GDCTX.L2.S1*

*AC. TC nº 137/2001, Proc. nº 78/2000*

*AC. TC nº 137/2002, Proc. nº 363/01.*

*AC.STJ, 20-09-2017, Proc. nº 1353/13.6GBABF.E1.S1*

*AC.TC nº 394/1989, Proc. n.° 93/88.*

*AC.TC.160/2010, Proc. n.º 834/09.*

*Al-Khawaja & Tahery v. U.K., Eur. Ct. H.R., App. No. 26766/05 & 22228/06, 37 (2011).*