Teixeira, Marco André Moreira de Sá Assunção,
1976-
Araújo, Marisa Isabel Almeida, 1981-

Vis-à-vis artificial intelligence, human rights, democracy, and the rule of law

http://hdl.handle.net/11067/6481
https://doi.org/10.34628/qa73-2h93

Metadados

Data de Publicação 2023

Resumo Artificial Intelligence (AI) is redesigning the status quo and transforming the world as we know it. The promises for a better and improved (as well as different) life are enormous. Although the downsides and risks related to AI and AI-based systems are also palpable. In fact, there are no doubts AI presents both benefits and risks at a human rights (HR) level, to democracy and the Rule of Law. Concerns about privacy, freedom, labour, health, equality, or non-discrimination, among others, are at...

A Inteligência Artificial (IA) está a redesenhar o status quo e a transformar o mundo como o conhecemos. As promessas de uma vida melhor e melhorada (e também diferente) são imensas. Não obstante as desvantagens e os riscos associados à IA e aos sistemas baseados em IA também são palpáveis. De facto, não temos dúvidas que a IA apresenta benefícios e riscos e em qualquer vertente o impacto projeta-se ao nível dos direitos humanos, bem como na democracia e no Estado de Direito. Preocupações com a ...

Tipo bookPart

4

VIS-À-VIS ARTIFICIAL INTELLIGENCE, HUMAN RIGHTS, DEMOCRACY, AND THE RULE OF LAW

VIS-À-VIS INTELIGÊNCIA ARTIFICIAL, DIREITOS HUMANOS, DEMOCRACIA E ESTADO DE DIREITO

Marco Teixeira* e Marisa Almeida Araújo**

https://doi.org/10.34628/qa73-2h93

«*Our future is a race between the growing power of our technology and the wisdom with which we use it.*»

(HAWKING, 2018)

* Head of Programme Office of UNODC in Mozambique and Malawi.
marco.teixeira@un.org

** Assistant Professor at Lusíada University, Faculty of Law. Researcher at Centre for Legal, Economic and Environmental Studies (CEJEA).
almeidaaraujo@por.ulusiada.pt

ABSTRACT: Artificial Intelligence (AI) is redesigning the *status quo* and transforming the world as we know it. The promises for a better and improved (as well as different) life are enormous. Although the downsides and risks related to AI and AI-based systems are also palpable. In fact, there are no doubts AI presents both benefits and risks at a human rights (HR) level, to democracy and the Rule of Law. Concerns about privacy, freedom, labour, health, equality, or non-discrimination, among others, are at the epicentre of the discussion and it creates a tangible tension between supporting innovation and promoting fairness. AI is also changing the way we access and interpret information, the way we decide, including as members of a politic society, and the way we participate in the democratic process. AI also plays a critical role in the way private and public institutions function, and it will come the moment (if not already) that it influences the way governments operate. Within this framework, this Chapter has the overarching aim to map the impacts AI and AI-based systems have on individuals and societies at a HR level and outline the contour of a HR-centered ethical approach to guarantee an Ethical AI, in its design, development, and deployment.

KEYWORDS: Artificial Intelligence; AI-based Systems; Ethics; Human Rights; International Human Rights Law; Democracy; Rule of Law.

RESUMO: A Inteligência Artificial (IA) está a redesenhar o *status quo* e a transformar o mundo como o conhecemos. As promessas de uma vida melhor e melhorada (e também diferente) são imensas. Não obstante as desvantagens e os riscos associados à IA e aos sistemas baseados em IA também são palpáveis. De facto, não temos dúvidas que a IA apresenta benefícios e riscos e em qualquer vertente o impacto projeta-se ao nível dos direitos humanos, bem como na democracia e no Estado de Direito. Preocupações com a privacidade, liberdade, trabalho, saúde, igualdade ou não discriminação estão no epicentro do debate, o que gera uma tensão tangível entre a inovação, e o apoio ao desenvolvimento, por um lado, e a promoção da equidade por outro. A IA está a mudar a forma como acedemos e interpretamos a informação, a forma como decidimos, incluindo como membros da sociedade política, e a forma como participamos no processo democrático. A IA também desempenha um papel vital na forma como as instituições, públicas e privadas, funcionam e influenciará (se já não o faz) a forma como os próprios governos operam. Nesta visão, este Capítulo visa sobretudo explorar e mapear o impacto da IA, e dos sistemas baseados em IA, nas pessoas e nas sociedades ao nível dos direitos humanos e delinear o contorno de uma abordagem ética centrada nos Direitos Humanos para garantir uma IA Ética, no seu design, desenvolvimento e implementação.

PALAVRAS-CHAVE: Inteligência Artificial; Sistemas baseados em Inteligência Artificial; Direitos Humanos; Ética; Direito Internacional dos Direitos Humanos; Democracia; Estado de Direito.

1. Introduction

AI [1] and AI-based systems are part of our day-to-day life and there is no turning back. AI makes our lives easier and, generally, better.

From our virtual personal assistants and self-driving vehicles to robots [2, 3, 4, 5, 6] (and robotic industry), to health care [7, 8, 9, 10], economy, health, privacy, information [11], military, and many other sectors, AI and related technologies are pushing the boundaries – for good and, in some levels, it could be for the worst – fueled by Big Data [12, 13].

However, what is AI (and AI-based systems)?

The answer is not as simple as we might think.

For instance, the European Commission's Communication on AI defines AI as «[…] *systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals*.» [14] AI include AutoML [15] and machine learning (ML) [16, 17]. Although, a few months later, the European Commission's High-Level Expert Group on Artificial Intelligence defined AI and ended up proposing a new updated definition, that

> *«[…] 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 analysing 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).»* [18]

On 2021, the EU Draft Act on AI, in its Article 3, point 1, defines AI system as a

> *«[…] software that is developed with one or more of the techniques and approaches listed in Annex I [19] 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.»* [20]

Although, on early March 2022, the European Committee of the Regions (ECR) has published amendments to the AI Act, including the definition itself of AI. The Amendment 2 refers that the AI systems «[…] *is an ongoing process that should take into account the context in which AI operates, keep pace with social developments in this field and not lose sight of the link between the ecosystem of excellence and the ecosystem of trust*» [21], reasoning that that AI requires and adaptive and evolving approach. Adding, on Amendment 30 that the list on Annex 1 should not be exhaustive «[…] *and it must be clear that it is based on the current scientific state of play.*» [22]

Saying this the Committee proposes the following definition (Amendment 16) of AI systems means:

> *«[…] software that is developed with one or more of the techniques and approaches listed (non-exhaustively) in Annex I, combined with social practices, identity am culture, and that can, for a given set of human-defined objectives, by observing its environment through collecting data, interpreting the collected structured or unstructured data, managing knowledge, or processing the information derived from these data, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;»* [23]

We can understand the importance of the definition since it sets the scope of the legislative frameworks however, the concept is volatile, and it mostly depends on the approach adopted and how flexible we want the definition to be.

Despite the approach adopted, [24] the fact is that *intelligence* is now (also) a(n artificial) process [25] and a digital revolution is in course. Although AI is expected to bring benefits, there are tangible implications at not only a human rights level, but also democracy and the rule of law. For that reason, a «human-centric approach» to AI is pledged by policy-makers [26], [27], [28].

2. Overview of the EU's policy approaches to AI

The EU, besides a soft law approach is also delivering a specific regulatory framework about AI [29]. It was already recognized the need to build a European strategy for AI.

In addition, the USA recognized the main purposes, among others, to ensure the US leadership in AI research and development lead the world in the development and use of trustworthy AI systems and prepare the USA workforce for the integration of AI systems across all sectors of economy and society [30].

The US law has some recent developments.

On February 3, 2022, it was reintroduced the bill «Algorithmic Accountability Act of 2022» – with some minor updates since the 2019 version – to the Senate [31] and the House of Representatives [32]. The Act

> *«[…] aims to hold organizations accountable for their use of algorithms and other automated systems that are involved in making critical decisions which affect the lives of individuals in the U.S. Among other requirements, the U.S. Act would mandate covered entities to conduct impact assessments of the automated systems they use and sell in accordance with regulations that would be set forth by the Federal Trade Commission ("FTC"). »* [33]

On April 10, 2018, 25 European countries signed a declaration of cooperation on AI. The Member States recognized that AI could solve key challenges, and Europe's competitiveness should be assured in the research and deployment of AI.

Although AI also brings new or renewed queries and legal and ethical questions that should and need to be addressed [34].

The European Commission, in its 2018 communication, aligned the position built on the declaration of cooperation, and recognized the EU «[…] *should be ahead of technological developments in AI and ensure they are swiftly taken up across its economy.*» [35], [36].

Ursula von der Leyen's political guidelines [37] also aligned to these goals and included a topic on an «Europe fit for the digital age», recognizing that «[d]*ata and AI are ingredients for innovation that can help us find solutions to societal challenges* […]» [38].

A shared commitment with the USA.

The EU-US Trade and Technology Council (TTC) affirm the goals to «[…] *coordinate approaches to key global technology, economic, and trade issues; and to deepen transatlantic trade and economic relations, basing policies on shared democratic values*». The EU and the USA are clearly committed in working together «[…] *to ensure that AI serves our societies and economies and that it is used in ways consistent with our common democratic values and human rights*.» [39]

The Commission's communications - «Shaping Europe's digital future» and the «A European Strategy for Data» - emphasize the goal that is wanted a «[…] *European society powered by digital solutions that are strongly rooted in our common values, and that enrich the lives of all of us* […]» [40] and «[…] *Europe aims to capture the benefits of better use of data, including greater productivity and competitive markets, but also improvements in health and well-being, environment, transparent governance and convenient public services.*» [41]

Processing data and training AI is also a major challenge and computer capacity is essential to take over the task, as recognized already in 2019 by the European Commission [42].

The European High-Performance Computer Joint Undertaking (EuroHPC JU) – a joint initiative between the EU, European countries, and private partners – are developing the next generation of world-leading supercomputing and data infrastructure in Europe, to improve Europe's scientific excellence and industrial strength, support economic digital transformation and ensure technological sovereignty [43].

In the White Paper on Artificial Intelligence [44], it is recognized that AI is a strategic technology and the «European approach for AI» is centred in promoting «[…] *innovation capacity in the area of AI while supporting the development and uptake of ethical and trustworthy AI across the EU economy*» [45].

The initiative was preceded by a public consultation which is most relevant to collect stakeholders' opinions [46]. Moreover, the accompanying report on the safety and liability implication of artificial intelligence, the internet of things and robotics [47].

On 2021, the European Commission run a public consultation (from October 2021 to January 2022) on the inception impact assessment about the Commission's plans about the initiative of Adapting liability rules to the digital

age and circular economy [48]. The Commission adoption is planned for the third quarter of 2022.

The above mentioned report emphasizes the challenges created in terms of liability of the new digital technologies and those «[…] *liability related challenges need to be addressed to ensure the same level of protection compared to victims of traditional technologies* […]»

The 2021 Communication from the European Commission addressed the necessary deployment of trustworthy AI as a pre-condition for Europe's future competitiveness and prosperity [49].

In addition, the European Parliament adopted several resolutions on AI, e.g., on the matters of ethics [50], civil liability [51], intellectual property [52], criminal law and AI use by the police and judicial authorities in criminal matters [53], and education, culture, and audiovisual sector [54].

The Parliament also set up a Special Committee on AI in a digital age (AIDA). The Committee is vested to analyse the future impact of AI in the digital age on the EU economy, investigate the contribution of AI to business value and economic growth, analyse the approach of third countries and to submit to Parliament's responsible standing committees an evaluation defining EU objectives in the medium-and long-term [55]. The Special Committee adopted its final recommendations in March 2022, and it will be put to vote by the full house in May 2022.

On 2019, the Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission released their Ethics Guidelines for a Trustworthy AI [56]. In the report, the group establishes three main components for a trustworthy AI.

It should be lawful, ethical, and robust.

All of the components are necessary itself but not sufficient «[i]*deally, all three components work in harmony and overlap in their operation.*» [57]

However, besides some soft law instruments adopted by the Commission, such as the above-mentioned Ethics Guidelines for Trustworthy AI [58], the legislative major milestone to the date was achieved in April 2021 [59]. The European Commission unveiled a regulatory framework on AI systems and the associated [60].

The AI Act clearly sets the goal to establish a legal and harmonized framework on the digital governance, set up in a risk-based approach [61], considering the fundamental rights of the EU [62], and the values of the Union [63], to foster investment

and innovation in AI, enhance governance and enforcement, and encourage the single European AI market.

A coordinated European approach to deal with AI is on its way and reveals a commitment to a key dossier, putting EU as a world-leader in AI matters and setting a tone to Member-States in their national AI strategies.

The principle set by the Commission is that «*AI should work for people and be the force for good in society*» [64], including the core values of the Union, summarized in the Pittsburgh Statement, such as the respect for human rights, environmental protection, the rule of law, non-discrimination, regulatory transparency, market-based commerce, and the freedom to innovate and to have innovations protected [65].

3. The next chapter: some remarks on AI challenges

We can summarize the AI challenges into five categories following Powers and Ganascia. «[C]*onceptual ambiguities, the estimation risks, implementing machine ethics, epistemic issues of scientific explanation and prediction, and oppositional versus systemic ethics approaches*». [66]

The regulatory efforts on AI, either on a soft law approach, but also the recent hard law regulatory efforts, many of which led by the EU, are aligned to address them.

Although the arena in which AI is moving is treacherous, but also highly seductive. Finding a balance from a precautionary approach to promoting evolution is not always easy, especially when dealing with confronting interests, different regulatory levels, social and economic asymmetries, and cultural differences, besides the impact in the future generations' way of life.

However, despite the differences and asymmetries, technological development brought to our lives many of the science fiction scenarios, that are not fictional anymore, and we entered a new era in the history of the world.

This new chapter of world history is presenting a new *humanoid species* that can parallel human beings [67], [68] at the *top of the food chain*. Bill Joy's cautionary essay talks about a new Pandora box [69] and how it threatens the human species [70].

The growth of *techne* assumes, in the words of Hans Jonas, ethical significance considering the central place it occupies in human purpose [71] like in no

other issues; anticipation has never been such an imperative, especially considering that we are also changing, in fact, the very nature of the human actions. Just have in mind, for instance, robots and their interaction with humans. Robots that can, not only, interact with humans but also cooperate in a truthful human-robot interaction [72]. With social robots, human-like decision-making skills are sought to increase their capability to deal with humans, but also our own empathy towards the machine considering the human resemblance about their attitudes and responses [73].

AI and AI systems will exceed human performance and we may well be at the down of the next chapter in world history: «*Artificial super-intelligence versus you and me.*» [74]

In a rather dark perspective professed by Irving John Good, the *deus ex machina* will kill its creator [75] in a similar perspective «*Nietzsche made in God's image turned and philosophically sought to kill man*» [76]. In fact, the potential of AI is, as Bostrom states, «[…] *vastly greater than that organic intelligence.*» [77]

Truth is that cognitive computing is *potentially* able to imitate or overcome human cognitive capacities, «[…] *on the basis of algorithms that embody 'machine learning' or even 'deep learning', using 'neural networks' that mimic the functioning of the human brain*» [78].

But if modelling the human brain is a task to be fulfilled, AI is built in a data-driven approach – based on artificial neural networks (ANNs) [79] – and, such systems, learn to perform tasks [80] considering examples (labelled data) without being programmed with any task-specific rules or models [81] and are capable to execute them without, or very little, human involvement [82]. For instance, autonomous vehicles are capable of sensing the environment they are in. Nevertheless, if so, on another layer, like any human driver, it may be confronted with moral dilemmas. Think about the typical example. An accident where it's not possible to avoid hitting or hurt someone. The algorithm will choose who is sacrificed. The decision is based on logic and determine how the robot should act. How about common sense, empathy? How about moral?

Although, let us confront with an ethical query

«*[A] robot is walking to the post office to post a letter. […] Suddenly a toddler chases a duck which hops into the stream. The toddler slips and falls into the water*

which is one meter deep. The toddler is in imminent danger of drowning. The robot is waterproof. [...]» [83]

What should the robot do? [84], [85] Post the letter or save the child? The answer is morally simple, but the robot does not *feel* empathy and, to solve this query it needs rules to make a decision, i.e., to walk away from its chore and feels the urgency of the situation and act accordingly.

If we take ethics to concern, the question is «[...] *what would it take to build an ethical AI that could make moral decisions*?» [86]

Although, assuming a moral decision would presuppose a pre-requisite; a recognition of a moral personhood. However, is it possible to recognize Artificial Moral Agents? In addition, these artificial agents would, if so, have rights and responsibilities? [87] Would we recognize a new, and we use Agamben's term provocatively, *homo sacer* [88]? And, eventually, we would/could recognize it as one of us, identifying an artificial morality or even a *sacrum* that is able or could surpasse human beings?

Considering this question, what kind of a moral status AI could have, if some? And, if it does, or it will might have, «*How would we know whether or not we succeeded in creating genuine Artificial Moral Intelligence (AMI)*?» [89]

On another layer, AI-guided genetic-engineered nanotechnology, and robotics (AI-GNR) presents itself as a technological transformative-revolution [90] including the transformation or, for some, an improvement, of the human body and human self [91]. Of ours and the self's of the future generations.

For instance, the genetic genome editing (GGE). GGE is a

«[...] *set of technologies, including a new tool based on the CRISPR/Cas9 mechanism discovered in Streptococcus pyogenes. The system can be engineered to facilitate the targeted modification of specific DNA sequences in the genomes of living cells»* [92].

We are at the dawn of the *posthuman* era [93].

In this matter, we are the responsible creators of the next generation genetic modified humans, the *GenRich* humans [94]. Although these mutations «[...] *can never subsequently remove and will have to hand down not just to her own children but to all subsequent descendants»* [95].

The co-existence of Naturals and Genrich, using Silver's terms, or the emergence of post-persons, as Nicholas Agar refers to, can create not only a stratification of human beings, in the opposite direction of the human dignity principle, but can also lead, as Agar refers to, to the entitlement of the post-persons to demand sacrifices from human mere persons, as we now sacrifice sentient nonpersons [96].

The *Homo Artificialis*, in a transhumanist [97] perspective, represents the «*Homo Sapiens Sapiens accède au gouvernement de sa propre constitution*» [98], a truthful liberation – for some – of the human species [99], [100] from the limits of the body and, in some ways, of the self.

Habermas has a profound reflection about the intervention in human creation and the control over human nature that can lead to «*changes* [in] *the overall structure of our moral experience*» and the understanding of the individual's self.

In a stratospheric dimension, followers of Singularity [101], AI will eventually «[…] *merge with human brains and become an all-seeing, all-powerful, super-intelligence. For true believers, computers will augment and extend our thoughts into a kind of "amortality."* [102]» [103]. Although, acting upon human beings, in particular in cognitive enhancements that Bostrom summarizes three conclusion about a strategic cognitive enhancement:

> «*(1) at least weak forms of super –intelligence are achievable by means of biotechnological enhancements;*
> *(2) the feasibility of cognitively enhanced human adds the plausibility that advanced forms of machine intelligence are feasible […]; and*
> *(3) […] the probable emergence of a generation of genetic enhanced populations – voters, inventors, scientists – with the magnitude of enhancement escalating rapidly […]*» [104].

The action has the *virtue* of pushing natural selection to produce high-quality individuals brings back to memory dark episodes of the world history [105].

Another set of examples, in a different level, that brings other considerations, is a cultural and intellectual dimension of AI.

When AI gets artistic, and the algorithms not only create outputs but create art. In 2016 the "next Rembrandt" was created [106]. The Microsoft Project examined the artistic Rembrandt DNA. To create the "next Rembrandt" the entire

collection was examined, and it was studied the contents of Rembrandt's paintings pixel by pixel. The database was used as for the new artistic creation.

In the same dimension, in 2019, AI created music. Not some music, Shubert's music. AI neural network finished the symphony of Schubert, *powered by Huaway AI* [107].

Besides the intellectual property (IP) rights [108], there is also another array of interrogations that the matter arises. How do we feel about a creative non-human intelligence *surpassing* the human creation? Or, as in these examples, how do we feel about the non-human intelligence beaten the creator's mortality? How do we justice creativity? What is the value of human vs. non-human creativity? In addition, what is the place of these artistic creations in our cultural heritage?

Another issue that raises concern is related to autonomous weapon systems. These systems «[…] *select and apply force to targets without human intervention.*» [109]

The International Committee of the Red Cross (ICRC) reinforces the risks associated to these systems, which have a license to kill with loss of human control and, mainly, without human judgment.

The ICRC specially addresses concerns about legal perspectives, but also, ethical and humanitarian [110]. Providing a moral analysis to this topic will lead concerns related to civilian causalities and threats to HR, instability, unpredictability, responsibility, but also, hacking and cybersecurity [111].

The examples are presented as mere illustrations about the two sides of AI. If AI has a strong potential to enhance individual rights, social benefits and promote economic growth, on the other side, it raises equally great and palpable concerns, especially at the human rights level.

4. Human rights, democracy, and the rule of law to the AI domain

A trinitarian commitment to human rights, democracy, and the rule of law is, in the words of Mattias Kumm, « […] the dogma of the constitutionalist faith» [112].

Starting with the same reference, Paul Nemitz questions how new technology must be shaped to support the maintenance of constitutional democracy [113]. The author starts from a holistic look on the reality of technology and the accumulation of *digital* power in the hands of the «frightful five» [114] that can lead to

«[…] rises in stock market valuations, and therefore wield economic power which does not only guarantee disproportionate access to legislators and governments, but also allows them to handout freely direct or indirect financial or in kind support in all areas of society relevant to opinion building in democracy: governments, legislators, civil society, political parties, schools and education, journalism and journalism education and — most importantly — science and research.» [115]

In democratic processes, the misuses of AI are not new. Just consider electoral manipulation narratives.

The Cambridge Analytica scandal is one of the most palpable examples. It exposed unlawful *«[…] harvest Facebook data from millions of voters in the United Kingdom, the United States, and elsewhere enable malign actors to engage in political micro-targeting through the use of AI-driven social media content distribution systems […]»* [116].

The scandal had worldwide proportions and it showed how is easily possible to interfere with democratic election processes and influence political decisions.

In fact, AI systems are much more capable to subvert information and condition our choices, jeopardise freedom of choice and our right to receive truthful information, without interference.

It is undeniable that, in every aspect of our lives, individually and as members of society, algorithms play, nowadays, a critical role including the information *selected* to the way we analyze or interpret data and, subsequently, the decisions we make [117].

Technology has also been used to blurry the lines of truth.

Considerer the issue of disinformation by fake news which can be potentialized by digital tools. The deepfake is a AI based technology that can be defined *«[…] within the intersection of technology and communication and/or visual representation as 'a technology that uses Artificial Intelligence to produce or edit contents of a video or an image […]»* [118]. An image is morphed in another image or video to produce content that does not exist. It is a product of deep learning, and the process employs Generative Adversarial Networks (GANs) that can produce a novel content. This content can be released and spread, especially on social media. To maximize the potential the *«Digital Influence Machine»* has the aim to make the message reach the *right* people [119].

We can understand immediately the implications of deepfake news around consent for instance, although truth is that repercussions can be even more profound. Just consider political deepfake for instance. If fake news, in general, have the ability to threat democracy itself, deepfake not only can be used to the same goal but also enhance the potential alarmingly. One of the clearest examples of the last years were the 2016 and 2020 US presidential elections.

A study of Hunt Allcott and Matthew Gentzkow about the 2016 American elections reveals that fake news «[…] *was both widely shared and heavily tilted in favor of Donald Trump.*» [120], [121] The authors do not claim or conclude that fake news have been pivotal in the 2016 US presidential election, but it was estimated that the average of US adult read, at least, one fake news article during election period with higher exposure to pro-Trump articles [122].

In the political deepfake arena, the hyperrealism poses a different level of menace. What to do if we cannot trust what we see anymore? Probably one of the most well-known examples of deepfake is the video of a drunk Nancy Pelosi [123] and, recently a deepfake video of the Ukrainian President, Volodymyr Zelensky, telling his country people to surrender [124].

Key decisions on AI development need a democratic debate and a multiple public engagement, from different stakeholders, from the North and South of the Globe, in a global dialogue, not only related to economic, social, and regulatory policies, but also diplomatic and political [125].

Truth is that the main issues in AI relate, among others, with accessibility, integrity, privacy, safety, bias, explainability and transparency.

These challenges were addressed in the coordinated plan on AI [126] and the 2021 review proposed key actions on how to create EU global leadership and trustworthy AI, such as accelerate investments in AI technology, act on AI strategies and programs, and align AI policy [127] to remove fragmentation [128].

The impact of AI and AI systems in people's lives and in societies in general has been calling for a HR approach and response to the increasing pressure it has been placing on human rights and the human rights discourse.

In sum, HR and fundamental freedoms that can be impacted by AI systems are, among others: human dignity, freedom, autonomy, fairness, non-discrimination, equality, diversity, privacy, data protection, democracy and the rule of law, besides cultural, social, and economic rights.

5. On the route for an ethical AI

The father of AI is recognized to be Alan Turing [129] that, in the 1950's published the well-known article «Computing Machine and Intelligence» and set the tone for the discussions we are still having nowadays. Alan Turing replaced the key question «*Can machines think?*» to the Turing Test (TT) [130].

The TT departs from and imitation game [131] to a new question: «*What will happen when a machine takes the part A of this game?*» [132]. Turing also recognized the difficulty of the learning machines and how the process should be similar to the education process of a child, making an

«*[…] obvious connection between this process and evolution, by the identifications*

Structure of the child machine	=	*Hereditary Material*
Changes of the child machine	=	*Mutations*
Natural selection	=	*Judgement of the experimenter*» [133]

Although, the learning process is, recognizably, complex, and how suitable it could be for a «child machine», mainly having in consideration that the pupil can be far more intelligent and have more knowledge than the professor, is one of the multiple questions that arise.

However, not being our goal to analyze Turing's work, it is impossible not to recognize the influence on the topic, and especially the mimic between humans and machinery evolution, and the multilateral connections between the two.

The machine learning process is particularly interesting at this level since – considering human learning – it is no one way process, there is a mutual flux, not only from the teacher to the student, but also vice versa [134].

In addition, a particular aspect of the work arises, and it's related to the correlation between language – and its multiple opacities – to cognitive abilities and knowledge. Moreover, knowledge is complex and «[t]*o create flexible intelligence in our machines we need to automate the knowledge-acquisition process.*» [135] To achieve that, in Kurzweil's words, it is imperative to create a way the system can model and understand human language and knowledge [136].

In a particularly interesting analyses of Turing's TT, Saygin, Cicekli and Akman conclude that it is about simulating human use of language by computers but, as they put it, it raises a new set of questions:

> «How do humans use language in similar settings? What is the relation between language and cognition? Is language autonomous with respect to other cognitive abilities? How can computers be made to understand language? What does a "simulation" mean, anyway?» [137]

The correlation between language and knowledge and the process of learning is rather important since, the progression from the human child and machine learning, follows reverse directions [138]. Children start listening and understanding spoken language and then written language, and machines start with written language and then understanding it, and only after listening and then understanding spoken language [139].

One of the most puzzling challenges is that, as Ammanath recalls, AI does not understand what it accomplishes since «[…] *unlike humans, AI lacks an internal model of the world*» [140].

Besides all of this, another variable needs to be considered.

The machine learning process needs to be built upon a set of principles. However, which ones? One of the most the most popular machine (pseudo-)ethics came from Isaac Asimov's Laws of Robotics in *I Robot*.

The three laws are:

> «*A robot may not injure a human being or, through inaction, allow a human being to come to harm.*
>
> *A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.*
>
> *A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.*» [141] *(Salge, 2017)*

Actually, there is another zeroth law added later by the author:

«0. *A robot may not harm humanity, or, by inaction, allow humanity to come to harm.*» [142], [143].

Friedman and Nissenbaum pioneered the process "[…] *of building systems for human-computer interaction that prioritize trust and user welfare while endeavoring to reduce biased outcomes.*" [144]

In a recent statement, [145] the United Nations' Human Rights Chief, Michelle Bachelet, expressed her concern about AI technologies. For instance, one of the most recent are Pegasus revelations spread suggesting a software abuse.

«*The Pegasus malware infects electronic devices, enabling operators of the tool to obtain messages, photos and emails, record calls, and even activate microphones, according to the consortium's reporting. The leak contains a list of more than 50,000 phone numbers which reportedly belong to those identified as people of interest, by clients of the company behind Pegasus, including some governments.*» [146]

According to the report of the Organized Crime and Corruption Reporting Project (OCCRP) corrupt regimes can gain access to personal information about anyone they want, and it is also reported that has been used against journalists, activists and political dissidents [147].

The United Nations' Human Rights Chief also addressed, in the same statement [148] the opaque deals of a different spyware, Candiru that emerged about the same time. Candiru «[…] *is a secretive Israel-based company that sells spyware exclusively to governments. Reportedly, their spyware can infect and monitor iPhones, Androids, Macs, PCs, and cloud accounts.*»

The report that targeted Candiru

«[…] *identified more than 750 websites linked to Candiru's spyware infrastructure.* […] *many domains masquerading as advocacy organizations such as Amnesty International, the Black Lives Matter movement, as well as media companies, and other civil-society themed entities.*» It was also identified «[…] *a politically active victim in Western Europe and recovered a copy of Candiru's Windows spyware.*» [149]

Although recognizing AI can be used to help societies to overcome challenges and improve people's lives, it has also a negative side and it can affect human rights, democracy, and the rule of law.

In the report of Thorbjørn Jagland to the 129th Session of the Committee of Ministers, new challenges to humankind were emphasised

«[...] for which Council of Europe legal standards are required. Three immediate challenges stand out: how to harness the benefits of the artificial intelligence revolution, while identifying and mitigating its threat to human rights, democracy and the rule of law» [150].

These risks are not a price to pay, and Michelle Bachelet calls for an urgent action to assess the risks, and concluded that «[…] *until compliance with human rights standards can be guaranteed, governments should implement a moratorium on the sale and transfer of surveillance technology.*» (Bachelet, 2021)

In fact, the examples teem.

For instance, ProPublica analysed a commercial AI tool made by Northpointe, Inc. and tested whether the recidivism algorithm [151], the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) used in the US judicial system, was predisposed against certain groups [152].

The analysis, reported on May 23, 2016 - «How we analysed the COMPAS Recidivism Algorithm» -, found that black defendants were more likely, than white ones, to be incorrectly judge to be at higher risk of re-offense [153]. The report concluded that:

«Black defendants were twice as likely as white defendants to be misclassified as a higher risk of violent recidivism, and white recidivists were misclassified as low risk 63.2 percent more often than black defendants. Black defendants who were classified as a higher risk of violent recidivism did recidivate at a slightly higher rate than white defendants (21 percent vs. 17 percent), and the likelihood ratio for white defendants was higher, 2.03, than for black defendants, 1.62.» [154]

The algorithms depend mostly on the data fed to them and, if so, they will reproduce and recreate bias and stereotypes [155], [156], [157] and, data fundamentalism [158] can represent a true threat to many people and impact their lives negatively.

Although, besides this particular level, another one needs to be considered. The confrontation with IP rights and transparency.

A significant case was discussed in the USA. In the case Wisconsin v. Loomis [159] the Wisconsin Supreme Court analyzed for the first time the use of algorithms and the right to a due process [160]. Loomis defense argued that, among others,

it was not possible to know and understand how the algorithms used predicted the defendant's recidivism since it was protected by IP rights. The defendant was unable to challenge the validity of the risk assessment produced by COMPAS tools.

The European Commission for the Efficiency of Justice (CEPEJ) in its European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment [161] addressed the specific issues to criminal justice, especially prevention of offences, risk of recidivism and assessment of the level of danger [162]. The CEPEJ addressed the potential risks of discrimination considering the AI Tools are

> «[…] constructed and interpreted by humans, can reproduce unjustified and already existing inequalities in the criminal justice system concerned; instead of correcting certain problematic policies, technology may end up legitimizing them.» [163]

The Charter defined five core principles in the field of Artificial Intelligence and justice: respect of fundamental rights, non-discrimination, quality and security, transparency, impartiality, and fairness, "under user control".

Also, the European Parliament resolution on AI in criminal law and its use by the police and judicial authorities in criminal matters

> «[…] recalls that the inclusion in AI training data sets of instances of racism by police forces in fulfilling their duties will inevitably lead to racist bias in AI-generated findings, scores, and recommendations; reiterates its call on Member States, therefore, to promote anti-discrimination policies and to develop national action plans against racism in the field of policing and the justice system; […]» [164]

The Parliament also underlines that many algorithmically driven identification technologies «[…] *disproportionately misidentify and misclassify and therefore cause harm to racialized people, individuals belonging to certain ethnic communities, LGBTI people, children and the elderly, as well as women*; […]» [165], [166].

Besides that, the lack of representativeness [167] can condemn many, and their specific issues, to a limbo [168].

For instance, Sacha Costanza-Chock explains, from a personal experience point of view, as a non-binary trans female, some issues not easily resolved by

algorithms [169] since they are encoded through the design of norms, values and assumptions that leaves (or can leave) many behind. The author explains that «[…] *the current path of AI development will produce systems that erase those of us on the margins* […]» [170].

A clear example is the automatic gender recognition (AGR) that assumes that «[…] *gender is a static concept that does not frequently change across time and cultures.*» [171] The simplistic way, to say the least, to present gender can have highly negative outcomes to transgender community.

In another perspective, but considering the same concern, i.e., the particular queries within the margin population, Petra Molnar debates how technology is used in the migration queries and how, in her perspective, the States explore new technologies but, mainly, the lack of international regulation, to deal with migration [172] .

The issues raised by race, class, and gender inequality through AI have been addressed and the author recalls Crenshaw's intersectional analyzes [173], as a powerful tool for the development of AI. A non-intersectional «[…] *algorithmic bias audits are insufficient to ensure algorithmic fairness.*» [174]

In AI systems we have some other examples that emerge. For instance, in 2018 Reuters reported that Amazon.com Inc's AMZN.O machine learning specialist uncovered that «[…] *their new recruiting engine did not like women.*» [175] The hiring automate tool reflected a male dominance so, therefore, the Amazon's AI systems learned that male candidates were more suitable than the female candidates [176].

Although, we might also consider another level of issues related to data. Synthetic data, which is generated artificially [177]. Some of the main challenges is how synthetic data is produced, how to ensure the characteristics between the original and synthetic data and what method and how to measure it [178].

The matter of bias and discrimination is specifically addressed in the Executive Office of the President's Report [179] that grouped the challenges to promote fairness and overcoming the discriminatory effects of data: challenges related to the data used as inputs and challenges related to the inner workings of the algorithm itself. In the first group, the report the decision to use certain data inputs can result in discriminatory outputs. Such as poorly selected data, incomplete, incorrect, or outdated data, selection bias, and unintentional perpetuation and promotion of historical biases.

On the second group, the report highlights the flaws related to poorly designed matching systems, personalization and recommendation services that narrow instead of expand user options, decision-making systems that assume correlation necessarily implies causation, data sets that lack information or disproportionally represent certain populations [180], and that algorithms encode discrimination and bias outputs related to the different participation in the digital ecosystem, «[…] *due to economic, linguistic, structural or socioeconomical barriers, among others.*» [181]

The 2019 Declaration adopted by the Committee of Ministers on the manipulative capabilities of algorithmic processes raised recalled about the fast-growing presence of technology in our lives. Although public in general is not aware the extent of data collection everyday devices gather and how «[t]*hese data are used to train machine-learning technologies to prioritise search results, to predict and shape personal preferences, to alter information flows, and, sometimes, to subject individuals to behavioural experimentation.*» [182]

It is undeniable that AI entrenches bias, and, for that reason, outputs can not only replicate but also amplify it and, even, be weaponized against certain groups of people. In another aspect, AI is mainly a product of the Global North and, if so, «*if privileged white men are designing the technology and business models of AI, how they design for the south?*» [183]

Therefore, not only the amount of data is crucial to mitigate discriminatory outputs but also, guarantee a diversity of data to cover most spectrums. And, besides the substantial and diverse data, as stated by the AI Act, it is an obligation of States for *ex ante* test, manage risk and guarantee human oversight and, also, to guarantee that AI systems maintain «[…] *its level of performance under any circumstances*» [184]. These obligations will facilitate the respect of other fundamental rights by minimising the risk of erroneous or biased AI-assisted decisions.

Promoting a fair and impartial AI embraces, as an imperative, diversity and inclusion.

On February 2022, the UN Human Rights Office of the High Commissioner released information about how proper data collection can change people's lives, «*in a people-centred approach to data collection*» referred by the UN Assistant Secretary-General for Human Rights, Ilze Brands Kehris. The UN Assistant Secretary-General referred specifically to the

«[…] lack of data, disaggregated by race or ethnic origin, as well as by gender, age, and other factors, hides the disproportionate impact of certain laws, policies and practices on racial or ethnic groups in all areas of life, from housing and education to employment, health and the criminal justice system.» [185]

On September 2019, the Council of Europe's (CoE) Committee of Ministers adopted the terms of reference for the Ad Hoc Committee on Artificial Intelligence (CAHAI), mandated

«[…] to examine, on the basis of broad multi-stakeholder consultations, the feasibility and potential elements of a legal framework for the development, design and application of artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law.» [186], [187]

In December 2020, CAHAI published a compilation of contributions [188] and adopted its Feasibility Study on a legal framework on AI design, development and application based on CoE standards [189], [190].

The Alan Turing Institute prepared a program to support the Feasibility Study published by the Council of Europe's Ad Hoc Committee on Artificial Intelligence. The Human Rights, Democracy and the Rule of Law Assurance Framework (HUDERAF) was completed in September 2021 and *«[…] combines the procedural requirements for principles-based human rights due diligence with the governance mechanisms needed to set up technical and socio-technical guardrails for responsible and trustworthy AI innovation practices»* [191].

The range of decisions that AI systems can take have – as the examples demonstrates – a direct impact at a HR level, especially from the most vulnerable ones.

The acknowledgement that AI has a significant impact on people's lives – individually, socially, and politically – urges not only policymakers, but also technological companies [192] and other stakeholders, including non-profit organizations and academia [193], to recognize the need for an AI Ethics. There are several examples of international efforts [194]. For instance, the Toronto Declaration led by Amnesty International and digital rights group AcessNow [195], the Montreal Declaration for a responsible AI was an initiative of the Université de Montréal [196], the Asimolar AI principles [197], the Barcelona Declaration [198], ethical guidelines

from the Japanese Society for Artificial Intelligence [199], or the UNI Global Union Top 10 Principles for Ethical AI [200].

At the EU level, on 2018, the European Group on Ethics in Science and New Technologies also published a relevant statement on AI, Robotics and 'Autonomous' Systems and proposed a set of basic principles and democratic prerequisites, based on fundamental values: human dignity, autonomy, responsibility, justice, equity and solidarity, democracy, rule of law and accountability, security, safety, bodily and mental integrity, data protection and privacy and sustainability [201].

On November 2021, UNESCO adopted the Recommendation of the Ethics of AI [202]. Among its values, it establishes the respect, protection and promotion of human rights and fundamental freedoms, and human dignity, ensuring diversity and inclusiveness, living in peaceful, just, and interconnected societies.

The core principles listed are proportionality and do no harm, safety and security, fairness and non-discrimination, sustainability, privacy and data protection, human oversight and determination, transparency and explainability, responsibility and accountability, awareness and literacy and multi-stakeholder and adaptive governance and collaboration.

Trust in technology is a key feature and the primary reason for acceptance. Building trust in technology – the digital trust – depends on several factors [203]. Keng Siau abd Weiyu Wang list them as human, technology, and environment characteristics [204]. Paying special attention to the environment characteristics, trust incorporates also a cultural dimension [205].

Huang and Bashir conducted a survey, and the results indicate a correlation between cultural values and trust [206]. For instance, Tae Wan Kim, Tong (Joy) Lu, Kyusong Lee, Zhaoqi Cheng, Yanhan Tang and John Hooker analysed the tolerability of lying and deception in Conversational Artificial Intelligence (CAI) and how a cultural-specific approach influenced the response to permissibility or impermissibility [207], [208]. The authors analyse the different manners in negation considering different cultures and different behavioural norms suggesting that if, for example

> «[…] the degree of invitation of trust in human-to-human price negotiation
> in a Scandinavian country is believed to be significantly higher than in the U.S.

The same thing can happen in Human-to-AI interaction. Then, a conversational AI ethically optimized for U.S. consumers is not ethically ideal for consumers in Scandinavian countries.» [209]

Such sectorial aspects should be taken into account especially considering the different AI domestic legal solutions that can (or already have) been adopted, moreover since an active policymaking on AI has recently emerged at the national level [210].

In addition, besides the socio-cultural level, a religious level should not be erased, and it should be particularly interesting to, also, address a conversation between AI and religion. A theological approach can contribute to a critical thinking considering, to begin with, the impact AI has on the conceptions of the human being, the personhood, the world [211].

The topic will need further study, but we can already say that AI will push religious boundaries and conceptions [212], on one side and, on another one, it assumes a stimulating critical perspective on AI [213] since religions integrate many of the challenges that AI arises [214].

Our aim is to highlight a necessary democratic domain in AI Ethics but, also, a diversity one, able to accommodate relevant cultural aspects and how they can point to different ethical approaches. We have to make clear that we are not defending a relativist approach, especially considering the human rights arena where the subject is inserted. Although, it is critical to accommodate social-cultural differences, as Stephen Cory Robinson concludes, «[…] *to understand how cultural values interact with policy discussions about technologies such as AI*» [215].

The Ethics Guidelines for a Trustworthy AI [216] brings together different stakeholders and works upon the European Group on Ethics in Science and New Technologies progress on the matter.

The Guidelines address the impact of AI on fundamental rights and establishes three components for a trustworthy AI:

«1. it should be lawful, complying with all applicable laws and regulations;
2. it should be ethical, ensuring adherence to ethical principles and values; and
3. it should be robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.» [217, 218]

The document sets key requirements for AI to be deemed trustworthy: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination, and fairness; societal and environmental well-being and accountability [219].

Although, if the most challenging quest is to align AI goals and human rights, it is also imperative to have in mind the sense of the big picture, where no one is left behind.

6. AI literacy for a democratic dialogue and a legal-ethical sustainable development

«AI is everyone's business!»

This is the motto of UNESCO to make AI part of our – everyone's – debates [220]. In order to comply to Article 27 of the Universal Declaration of Human Rights it is essential to ensure everyone's participation, promote diversity, including the data collection and classification [221], [222], and access to technology.

Public awareness and literacy for AI is an imperative.

Understanding the different nuances and consequences is the only route to ensure that everyone single person can and is able to take an active role in the democratic dialogue about AI and is able to take informed decisions about the use of AI systems and the consequences of its use. As stated by UNESCO

> *«Learning about the impact of AI systems should include learning about, through and for human rights and fundamental freedoms, meaning that the approach and understanding of AI systems should be grounded by their impact on human rights and access to rights, as well as on the environment and ecosystems.»* [223]

Public awareness about social good can influence corporate social responsibility (CSR), that can include the ethics of AI that can help «[…] *bring together the notions of the social contract and responsible AI use*.» [224]

In fact, public must be prepared for AI, and this demands that education is required. Also, education depends on some variables, like the recipients, the speaker, and the context.

A public empowerment through education [225], to participate in decisional processes that are critical to shape our future. The HR approach to an ethical AI centrally claims self-determination through the means of democracy.

In 2016, the European Commission adopted a plan to enhance education related to labour market. The «New skills Agenda for Europe»[226], [227] and the «Digital Education action Plan» [228].

In fact, «technological unemployment» is a main concern. Like the industrial revolution, the labour force and labour market are changing [229].

A model of «*Explainable Artificial Intelligence (XAI)*», in the words of Martin Ebers, is not only relevant to the professionals [230], but also to all of those that might be affected by it [231]. Truth is that explainability demands for educated people, not only capable to understand the information, but also make critical assessments about it.

The «digital divide» [232] can exacerbate inequalities, bias, and exclusion and it is crucial to continue to question and test the inputs and the results they produce, the algorithm «black box» [233].

Transparency in AI is also a key element in AI systems. The OECD Recommendation recalls that, not only AI actors, should commit to transparency and disclosure but also to «[…] *provide meaningful information, appropriate to the context* […]» [234], the burden and responsibility to assure that recipients fully understand and interpret information. To achieve transparency AI system's decisions has to be traceable, explainable, and whether the AI system's capabilities and limitations have been communicated to recipients [235].

The Committee of Ministers encouraged Member States to promote digital literacy and enhance public awareness, «[s]*pecifically, public awareness should be enhanced of the fact that algorithmic tools are widely used for commercial purposes and, increasingly, for political reasons, as well as for ambitions of anti- or undemocratic power gain, warfare, or to inflict direct harm*;» [236]

In addition, domestic plans are considered a key aspect to address the issue to empower people through education and skills. Von der Leyen is committed to make the European Education Area a reality by 2025, bringing down barriers to learning and improve access to quality education, aligning the efforts to the Sustainable Development Goals of the UN also in AI area [237], [238].

7. Conclusions

This Chapter has highlighted the tension between AI and, HR, Democracy, and the Rule of Law, and the need to deliver an Ethical AI.

Identify the core values to establish ethical standards is imperative, and a substantial input from the public in general, and diverse stakeholders, guarantees a most-needed democratic dialogue and, also, that no one is left behind.

AI is up-to-date largely unregulated and a claim to a legislative movement is necessary. Like the new general data protection regulation (GDPR) [239] is a clear example of the relevance of changes and legislative initiatives that comes with progress of technology towards the public interest [240].

Complex technological innovations developed and applicated under opaque procedures difficult evaluation and assessment, and, with limited control, comes limited accountability. Not only in the outcomes produced but also in the AI architecture itself. Legislative initiatives to regulate AI systems – like the EU AI Act [241] – support the need for a homogenous legal regime that can address the multiple policy goals in order to protect people's rights, especially our fundamental rights and freedoms [242].

In addition, it will also have to strengthen democracy and the rule of law, providing the necessary legal tools to combat the resistance to compliance and accountability to adequately protect those rights. It is especially important considering that technology and AI systems are a global phenomenon and do not recognize the physical frontiers of Nations and States.

Oversight, compliance, and accountability procedures are – within a legislative framework, based on ethical principles – imperative to protect the triad human rights, democracy, and the rule of law in the challenging AI arena.

8. A note on methodology

A study about an interdisciplinary and broad topic like the one presented in this Chapter is always a difficult academic quest. This study is fundamentally concerned on the ethical and legal implications of AI and AI-based systems and how human rights, democracy and the rule of law are affected by it. With this objective in mind, the study used qualitative methods, including legal, ethical and

philosophical analyses, resorting to the work developed by international actors, including ongoing and work in progress projects, over the topic and a review of literature related to the subject of this study.

9. Acknowledgements

10. Competing interests

The authors declare they have no competing interests.

11. Funding

References

Ad Hoc Committee on Artificial Intelligence. (2020, December 17). *Feasibility Study.* Retrieved April 14, 2022, from https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da

Ad Hoc Committee on Artificial Intelligence. (2020, September 23). *Progress report.* Retrieved March 17, 2022, from https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016809ed062

Agamben, G. (1998 (originally published in 1995)). *Homo Sacer: Sovereign Power and Bare Life* (2nd ed.). (D. Heller-Roazen, Trans.) USA: Stanford University Press.

Agar, N. (2014). *Truly Human Enhancement. A Philosophical Defense of Limits.* Massachusetts: MIT Press.

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives, 31 (2)*, 211-236. doi=10.1257/jep.31.2.211

Almeida Araújo, M., & Meireis, A. (2019). Understanding Dignity (of Workers) across the Catholic Thought. *Minerva: Revista de Estudos Laborais*, 27-62. Retrieved from https://www.researchgate.net/publication/339845373_UNDERSTANDING_DIGNITY_OF_WORKERS_ACROSS_THE_CATHOLIC_THOUGHT

Almeida Araújo, M., & Meireis, A. (2020). A Century After Plessy v. Ferguson. The Struggle for Racial Equality in the Algorithm Era. *Lusiada Law Review, 23/24*, pp. 195-204.

Ammanath, B. (2022). *Trustworthy AI: A Business Guide for Navigating Trust and Ethics in AI.* New Jersey: Wiley.

Artificial Intelligence Index Report 2019. (2019). *Stanford University.* Retrieved May 1, 2022, from https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf

Arun, C. (2020). AI and the Global South: Designing for Other Worlds. In M. D. Dubber, F. Pasquale, & S. Das, *The Oxford Handbook of Ethics of AI* (pp. 588-607). Oxford: Oxford University Press.

Asaro, P. (2020). Autonomous Weapons and the Ethics of Artificial Intelligence. In S. M. Liao, *Ethics of artificial Intelligence* (pp. 212-236). Oxford: Oxford University Press.

Bachelet, M. (2021, September 14). Committee on Legal Affairs and Human Rights, Parliamentary Assembly Council of Europe. Retrieved January 3, 2022, from https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=27455&LangID=E

Barrat, J. (2013). *Our Final Invention: Artificial intelligence and the end of the human Era.* New York: St. Martin's Press.

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An Introduction to Ethics in Robotics and AI.* Springer.

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society , 20:1*, 1-13. doi: https://doi.org/10.1080/1369118X.2016.1216147

Ben-Israel, I., Cerdio, J., Ema, A., Friedman, L., Ienca, M., Mantelero, A., . . . Vayena, E. (2020, December). *Towards a 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.* Retrieved April 10, 2022, from Compilation of contributions prepared by the CAHAI Secretariat: https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a

Bostrom, N. (2003). *The Transhumanist FAQ - A General Introduction - Version 2.1 .* Retrieved September 9, 2020, from World Transhumanist Association: https://www.nickbostrom.com/views/transhumanist.pdf

Bostrom, N. (2017 (first published in 2014)). *Superintelligence: Paths, Dangers, Strategies.* Oxford: Oxford Unversity Press.

Bryson, J. J. (2020). The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law and Regulation. In M. D. Dubber, F. Pasquale, & S. Das, *The Oxford Handbook of Ethics of AI* (pp. 2-25). Oxford: Oxford University Press.

Buchanan, A. (2011). *Beyond humanity?* Oxford: Oxford University Press.

Committee of Ministers. (2019, February 13). Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes.

Decl(13/02/2019)1. Retrieved February 16, 2022, from https://search.coe. int/cm/pages/result_details.aspx?ObjectId=090000168092dd4b

Committee of Ministers. (2021, March 17). Declaration by the Committee of Ministers on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of the social safety net. *Decl(17/03/2021)2*. Retrieved February 16, 2022, from https://search.coe. int/cm/Pages/result_details.aspx?ObjectId=0900001680a1cb98

Costanza-Chock, S. (2018, July 18). Design Justice, A.I., and Escape from the Matrix of Domination. *Journal of Design and Science*. doi:10.21428/96c8d426

Crenshaw, K. W. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review, 43(6)*, 1241-1299.

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data, 6:54*, 1-25. https://doi.org/10.1186/ s40537-019-0217-0

Declaration of cooperation on Artificial Intelligence (AI). (2018, April 10). Retrieved February 12, 2022, from https://wayback.archive-it. org/12090/20210104133957/https://ec.europa.eu/digital-single-market/ en/news/eu-member-states-sign-cooperate-artificial-intelligence

Donath, J. (2020). Ethical Issues in our Relationship with Artificial Entities. In M. D. Dubber, & S. D. Frank Pasquale, *The Oxford Handbook of Ethics of AI* (pp. 52-73). Oxford: Oxforxd University Press.

Ebers, M. (2021, July 15). Regulating Explainable AI in the European Union. An Overview of the Current Legal Framework(s). In L. Colonna, & S. Greenstein, *Nordic Yearbook of Law and Informatics 2020: Law in the Era of Artificial Intelligence.* Retrieved May 1, 2022, from in: Liane Colonna/ Stanley Greenstein (eds.), Nordic Yearbook of Law and Informatics: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3901732

European Comission's High-Level Expert Group on Artificial Intelligence. (2018, December 18). Retrieved February 14, 2022, from https:// ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_ december_1.pdf

European Commission. (2021). *Digital Education Action Plan (2021-2027).* Retrieved April 18, 2022, from https://education.ec.europa.eu/focus-topics/digital-education/about/digital-education-action-plan

European Commission. (2016, June 10). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions a New Skills Agenda for Europe (COM(2016) 381 final).* Retrieved April 14, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016DC0381

European Commission. (2018, April 25). *Communication from the Commission on Artificial Intelligence for Europe COM(2018) 237 final.* Retrieved February 14, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN#footnote12

European Commission. (2019, April 8). *Building Trust in Human-Centric Artificial Intelligence (COM(2019) 168 final).* Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52019DC0168&qid=1620297270190

European Commission. (2020, February 19). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committe of the Regions Shaping Europe's digital future (COM/2020/67 final).* Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0067

European Commission. (2020, February 19). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions A European strategy for data (COM/2020/66 final).* Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066

European Commission. (2020, February 19). *Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics COM/2020/64 final.* Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1593079180383&uri=CELEX%3A52020DC0064

European Commission. (2020, February 19). *White Paper on*

       Artificial Intelligence - a European approach to excellence and trust COM(2020) 65 final. Retrieved February 15, 2022, from https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Commission. (2021, June 30). *Adapting liability rules to the digital age and circular economy.* Retrieved March 25, 2022, from file:///C:/Users/almeidaaraujo/Downloads/090166e5df292bd3.pdf

European Commission. (2021, December 13). Big Data. Retrieved February 15, 2022, from https://digital-strategy.ec.europa.eu/en/policies/big-data

European Commission. (2021, April 21). *Communication from the Commission to the European Parliament, the council, the European Ecnomic and Social Committee and the Committee of Regions: Fostering a European approach to Artificial Intelligence (COM(2021) 205 final).* Retrieved April 2, 2022, from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM%3A2021%3A205%3AFIN

European Commission. (2021). *Coordinated Plan on Artificial Intelligence 2021 Review.* Retrieved April 2, 2022, from https://digital-strategy.ec.europa.eu/pt/node/9757

European Commission. (2021). *Digital Education Action Plan (2021-2027).* Retrieved April 7, 2022, from https://education.ec.europa.eu/focus-topics/digital-education/about/digital-education-action-plan

European Commission. (2021, April 21). *Proposal for a Regulation on the European Parliament and the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial intelligence Act) and Amending Certain Union Legislative Acts COM(2021) 206 final.* Retrieved December 18, 2021, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

European Commission for the Efficiency of Justice. (2018, December 3-4). *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment.* Retrieved April 3, 2022, from https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c

European Group on Ethics in Science and New Technologies. (2018, March 9). *Statement on Artificial Intelligence, Robotics and*

'Autonomous' Systems. Retrieved March 22, 2022, from Research and Innovation: https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1

European High Performance Computing Joint Undertaking. (n.d.). Retrieved March 26, 2022, from https://eurohpc-ju.europa.eu/discover-eurohpc-ju

European Parliament. (2020, June 18). *Decision of 18 June 2020 on setting up a special committee on artificial intelligence in a digital age, and defining its responsibilities, numerical strength and term of office (2020/2684(RSO)).* Retrieved January 6, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0162_EN.html

European Parliament. (2020, October 20). *Resolution on a civil liability regime for artificial intelligence, (2020/2014(INL)).* Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html

European Parliament. (2020, October 20). *Resolution on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)).* Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html

European Parliament. (2020, October 20). *Resolution on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)).* Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_EN.html

European Parliament. (2021, October 6). *Artificial intelligence in criminal law and its use by the police and judicial (2020/2016(INI)).* Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.pdf

European Parliament. (2021, November 17). *Briefing.* Retrieved February 14, 2022, from Artificial intelligence act: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792

European Parliament. (2021, October 6). *Resolution on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)).* Retrieved April 15, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html

European Parliament. (2021, May 19). *Resolution on artificial intelligence in education, culture and the audiovisual sector (2020/2017(INI)).* Retrieved January 7, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_EN.html

European Union Agency for Fundamental Rights. (2020). *Getting the Future Right. Artificial Intelligence and Fundamental Rights.* Retrieved April 16, 2022, from https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf

EU-US Trade and Technology Council Inaugural Joint Statement. (2021, September 29). Retrieved March 8, 2022, from https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_21_4951

Executive Office of the President. (2016, May). *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights.* Retrieved March 25, 2022, from https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

Figueiredo, E. A. (2021). *Direito e Nanobiotecnociência. Reflexões na encruzilhada da inovação, do risco e da crise do(s) direito(s).* Coimbra: Almedina.

Franke, U. (2021, June). *Artificial Intelligence diplomacy: Artificial Intelligencegovernance as a new European Union external policy tool.* Retrieved April 28, 2022, from Study Requested by the AIDA committee: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662926/IPOL_STU(2021)662926_EN.pdf

Freeman, K. (2016, 12). Algorithmic Injustice: How The Wisconsin Supreme Court Failed To Protect Due Process Rights in State v. Loomis. *North Carolina Journal of Law & Technology, 18*, pp. 75-106.

Fukuyama, F. (2002). *Our Posthuman Future. Consequences of the biotechnology revolution.* New York, NY: Picador.

Gebru, T. (2020). Race and Gender: Data-Driven Claims about Race and Gender Perpetuated the Negative Biases of the Day. In M. D. dubber, F. Pasquale, & S. Das, *The Oxford Handbook of ethics of AI* (pp. 252-269). Oxford: Oxford University Press.

Gellers, J. C. (2021). *Rights for Robots: Artificial Intelligence, Animal and Environmental Law.* New York: Routledge.

Greene, D., Hoffmann, A. L., & Stark, L. (2019). Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2122-2131. Retrieved April 13, 2022, from https://hdl.handle.net/10125/59651

Habermas, J. (2008). *The Future of Human Nature.* Cambridge: Polity Press.

Harari, Y. (2017). *Homo Deus: A Brief History of Tomorrow.* Random House.

Hawking, S. (2018). *Brief Answers to the Big Questions.* New York: Bantam Books.

Hood, L., Heath, J. R., Phelps, M. E., & Lin, B. (2004, October 22). Systems Biology and New Technologies Enable Predictive and Preventative Medicine. *Science, 306, Issue 5696*, pp. 640-643. DOI: 10.1126/science.1104635

Huang, H.-Y., & Bashir, M. (n.d.). Users' Trust in Automation: A Cultural Perspective. (J. Chen, Ed.) *Advances in Intelligent Systems and Computing , 595*, 282-289. doi: https://doi.org/10.1007/978-3-319-60384-1_27

Hutson, M. (2017, July 7). AI Glossary: Artificial intelligence, in so many words. *Science, 357, Issue 6346*, https://www.science.org/doi/10.1126/science.357.6346.19

Independent High-Level Expert Group on Artificial Intelligence to the European Commission. (2019, April 8). Ethics Guidelines for a Trustworthy AI. Brussels. Retrieved February 15, 2022, from https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf

Independent High-Level Expert Group on Artificial Intelligence to the European Commission. (2019, June 26). *Policy and Investment Recommendations for Trustworthy AI.* Retrieved February 15, 2022

Independent High-Level Expert Group on Artificial Intelligence. (2020, July 23). *Sectoral Considerations on the Policy and Investement Recommendations for Trustworthy Artificial Intelligence.* Retrieved February 16, 2022, from https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai

Independent Hugh-Level Expert Group on Artificial Intelligence. (2020, July 17). *The Assessment List for Trustworthy Artificial*

Intelligence for self assessment. Retrieved February 16, 2022,
from https://digital-strategy.ec.europa.eu/en/library/
assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

International Committee of the Red Cross. (2021, May 12). *ICRC position on autonomous weapon systems.* Retrieved May 3, 2022, from https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems

International Federation on Robotics. (n.d.). Topics and Definitions. Retrieved February 16, 2022, from https://ifr.org/#topics

International Telecommunication Union. (2012, June 15). *Overview of the Internet of things. Recommendation ITU-T Y.2060.* Retrieved March 23, 2022, from https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=y.2060

Ito, J. (2018, December 2). Resisting Reduction: A Manifesto. *Journal of Design and Science.* doi:10.21428/8f7503e4

Jagland, T. (2019, May 16-17). Ready for future challenges – reinforcing the Council of Europe. Helsinki. Retrieved February 14, 2020, from https://rm.coe.int/168093af03

James, A. (2020). Planning for Mass Unemployment: Precautionary Basic Income. In S. M. Liao, *Ethics of Artificial Intteligence* (pp. 183-211). Oxford: Oxford University Press.

James, S., Harbron, C., Branson, J., & Sundler, M. (2021, December 13). Synthetic data use: exploring use cases to optimise data utility. *Discover Artifcial Intelligence, 1(15).* https://doi.org/10.1007/s44163-021-00016-y

Joint Research Centre. (2019). *1st Workshop on Artificial Intelligence at JRC, Ispra 23rd May 2018.* doi:10.2760/705074

Joint Research Centre; European Commission. (2022, March 28). *AI Watch Index 2021.* Retrieved April 18, 2022, from https://ai-watch.ec.europa.eu/publications/ai-watch-index-2021_en

Jonas, H. (1984). *The Imperative of Responsability: In Search of an Ethics for the Technological Age.* Chicago: The University of Chicago Press.

Joy, B. (2000, April 1). *Why the Future Doesn't Need Us.* Retrieved April 20, 2022, from Wired: https://www.wired.com/2000/04/joy-2/

Kalpokas, I., Kalpokiene, J. (2022). Deepfakes: A Realistic Assessment of Potentials, Risks, and Policy Regulation, Switzerland: Springer.

Kim, T. W., Lu, T., Lee, K., Cheng, Z., Tang, Y., & Hooker, J. (2021, March 16). When is it permissible for artificial intelligence to lie?: A trust-based approach. *Ethics of Conversational AI*. Retrieved April 18, 2022, from https://arxiv.org/pdf/2103.05434.pdf

Kumm, M. (2013). The Cosmopolitan Turn in Constitutionalism: An Integrated Conception of Public Law. *Indiana Journal of Global Legal Studies, 20:2*, 605-628. Retrieved April 14, 2022, from https://www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=1516&context=ijgls

Kurzweil, R. (2000). *The Age of Spiritual Machines: When Computers Exceed Human Intelligence* . New York: Penguin books.

Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016, May 23). How We Analyzed the COMPAS Recidivism Algorithm: . *Propublica.* Retrieved February 16, 2022, from https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

Leslie, D., Burr, C., Aitken, M., Cowls, J., Katell, M., & Briggs, M. (2021). *Artificial Intelligence, human rights, Democracy, and the rule of Law: a primer. The Council of Europe.* Retrieved April 3, 2022, from https://rm.coe.int/primer-en-new-cover-pages-coe-english-compressed-2754-7186-0228-v-1/1680a2fd4a

Leslie, D., Burr, C., Aitken, M., Katell, M., Briggs, M., & Rincon, C. (2021). *Human rights, democracy, and the rule of law assurance framework for AI systems: A proposal. the Alan Turing Institute.* doi:https://doi.org/10.5281/zenodo.5981676

Leyen, U. v. (2019). *Political Guidelines for the next European Commission 2019-2024.* Retrieved March 9, 2022, from https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_en_0.pdf

Liao, S. M. (2020). The Moral Status and Rights of Artificial Intelligence. In S. M. Liao, *Ethics of Artificial Intelligence* (pp. 480-503). Oxford: Oxford University Press.

Marczak, B. B., Scott-Railton, J., Berdan, K., Razzak, B. A., & Deibert, R. (2022, January 3). Hooking Candiru: Another Mercenary Spyware Vendor Comes into Focus. Retrieved July 15, 2021, from https://citizenlab.ca/2021/07/hooking-candiru-another-mercenary-spyware-vendor-comes-into-focus/

Molnar, P. (2019, December). Technology on the margins: AI and global migration management from a human rights perspective. *Cambridge International Law Journal*, 305-330. doi:https://doi.org/10.4337/cilj.2019.02.07

Monlezun, D. J. (2020). *The Global Bioethics of Artificial Intelligence and Human Rights.* Cambridge: Cambridge Schoolar Publishing.

Moreira, A. J. (2021). *Escravidão, Dignidade e Trabalho.* Coimbra: Almedina.

Morgan, C. S., Langlois, F., & Lan, J. (2022, April 8). *U.S. House and Senate Reintroduce the Algorithmic Accountability Act Intended to Regulate AI.* Retrieved April 12, 2022, from Lexology: https://www.lexology.com/library/detail.aspx?g=c9bfbfca-a88e-4fd9-b556-d903a4cf0bd7

Müller-Abdelrazeq, S. L., Schönefeld, K., Haberstroh, M., & Hees, F. (2019). Interacting with Collaborative Robots—A Study on Attitudes and Acceptance in Industrial Contexts. In O. Korn, *Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction* (pp. 101-117). Springer.

Nadler, A., Crain, M. & Donavan, J. (2018). *Weaponizing the digital influence machine: The Political Perils of Online Ad Tech.* Data & Society Research Institute, https://datasociety.net/library/weaponizing-the-digital-influence-machine//

National Artificial Intelligence Initiative Act of 2020. (2022, March 3). Retrieved from https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210

Nedelkoska, L., & Quintini, G. (2018). Automation, skills use and training. *OECD Social, Employment and Migration Working Papers, 202.* doi:https://doi.org/10.1787/2e2f4eea-en

Nemitz, P. (2018, August 14). Constitutional democracy and technology in the age of artificial intelligence. *Phil. Trans. R.* doi:http://dx.doi.org/10.1098/rsta.2018.0089

OCCRP. (2021). The Pegasus Project. Retrieved May 15, 2022, from https://www.occrp.org/en/the-pegasus-project/

Odusami, M., Maskeliūnas, R., Damaševičius, R., & Krilavičius, T. (2021, June 10). Analysis of Features of Alzheimer's Disease: Detection of Early Stage from Functional Brain Changes in Magnetic Resonance Images

Using a Finetuned ResNet18 Network. *Diagnostics, 11 (6).*
https://doi.org/10.3390/diagnostics11061071

OECD Council. (2019, May 22). *Recommendation on Artificial Intelligence.*
Retrieved January 14, 2022, from https://legalinstruments.oecd.org/en/
instruments/OECD-LEGAL-0449

Opinion of the European Committee of the Regions. (2022, February 28).
*European approach to artificial intelligence - Artificial Intelligence Act
(revised opinion) (2022/C97/12).* Retrieved March 12, 2022, from Official
Journal of the European Union (C97/602).

Ormond, K. E., Mortlock, D. P., Scholes, D. T., Bombard, Y., Brody, L. C.,
Faucett, W. A., . . . Young, C. E. (2017, August 3). Human Germline
Genome Editing. *The American Journal of Human Genetics, 101 (2)*,
pp. 167-176. https://doi.org/10.1016/j.ajhg.2017.06.012

Policy Department for "Citizens' Rights and Constitutional Affairs. (2016,
October). European Civil Law Rules in Robotics. Retrieved February 17,
2022, from http://www.europarl.europa.eu/committees/fr/supporting-
analyses-search.html

Powers, T. M., & Ganascia, J.-G. (2020). The Ethics of the Ethics AI.
In M. D. Dubber, F. Pasquale, & S. Das, *The Oxford Handbook
of Ethics of AI* (pp. 26-51). Oxford: Oxford University Press.

Qlik. (n.d.). *AutoML. What it is, why you need it, and best practices.* Retrieved
February 20, 2022, from https://www.qlik.com/us/augmented-analytics/
automl

Qlik. (n.d.). *Big Data AI. How Big Data and AI Work Together.* Retrieved February
20, 2022, from https://www.qlik.com/us/augmented-analytics/big-data-ai

Raposo, V. L. (2022, March 18). Ex machina: preliminary critical assessment
of the European Draft Act on artificial intelligence. *International
Journal of Law and Information Technology, 30 (1)*, pp. 88-109.
doi: https://doi.org/10.1093/ijlit/eaac007

Reed, R. (2021, May 31). A.I. in Religion, A.I. for Religion, A.I. and Religion:
Towards a Theory of Religious Studies and Artificial Intelligence.
Religions, 12. doi: https://doi.org/10.3390/rel12060401

Robinson, S. C. (2020, November). Trust, transparency, and openness: How
inclusion of cultural values shapes Nordic national public policy strategies

for artificial intelligence (AI) . *Technology in society, 63*. doi: https://doi. org/10.1016/j.techsoc.2020.101421

Salge, C. (2017, July 11). Asimov's Laws Won't Stop Robots from Harming Humans, So We've Developed a Better Solution. *Scientific American*. Retrieved February 16, 2022, from https://www.scientificamerican.com/ article/asimovs-laws-wont-stop-robots-from-harming-humans-so-weve-developed-a-better-solution/

Savage, N. (2019, August 21). How AI and neuroscience drive each other forwards. *Nature*. Retrieved February 22, 2022, from https://www.nature. com/articles/d41586-019-02212-4

Saygin, A. P., Cicekli, I., & Akman, V. (2000, October). Turing Test: 50 Years Later. *Minds and Machines, 10(4)*, 463-518. doi:10.1023/A:1011288000451

Science and Technology Directorate. (2022, February 1). *Feature Article: Robot Dogs Take Another Step Towards Deployment at the Border.* Retrieved February 20, 2022, from https:// www.dhs.gov/science-and-technology/news/2022/02/01/ feature-article-robot-dogs-take-another-step-towards-deployment

Sève, L. (2006). *Qu'est-ce que la personne humaine?* Paris: La Dispute.

Siau, K., & Wang, W. (2018). Building Trust in Artificial Intelligence,. *Cutter Business Technology Journal, 31(2)*, 47-53. Retrieved April 16, 2022, from https://www.researchgate.net/profile/Keng-Siau-2/publication/ 324006061_Building_Trust_in_Artificial_Intelligence_Machine_ Learning_and_Robotics/links/5ab8744baca2722b97cf9d33/Building-Trust-in-Artificial-Intelligence-Machine-Learning-and-Robotics.pdf

Silver, M. L. (2007 (First Published 1997)). *Remaking Eden: How Genetic Engineering and Cloning will Transform The American Family.* USA: Ecco.

Singler, B. (2017). An Introduction to Artificial Intelligence and Religion For the Religious Studies Scholar. *The Faraday Institute for Science and Religion*, 215-231. doi:10.1558/imre.35901

Sirithunge, H. P., Bandara, H. M., Jayasekara, A. G., Chandima, D. P., & Abeykoon, H. M. (2019). A Study on Robot-Initiated Interaction: Toward Virtual Social Behavior. In O. Korn, *Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction* (pp. 33-70). Springer.

State of Wisconsin v. Loomis. Retrieved March 21, 2022, from https://casetext.com/pdf-email?slug=state-v-loomis-22

Terence, M. (2016). *Le Transhumanisme est un Intégrisme.* Paris: Les Éditions du Cerf.

Turing, A. (1950, October). Computing Machinery and Intelligence. *Mind. A Quarterly Review of Psycology and Philosophy, LIX (236)*, 433-460. Retrieved April 20, 2022, from https://watermark.silverchair.com/lix-236-433.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAuYwggLiBgkqhkiG9w0BBwagggLTMIICzwIBADCCAsgGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQMvKEkFW4xrg57Pl5CAgEQgIICmdlXdLEuv-7V6ssL2ndgapYY5NCh343GhNtuWUDTkhx

UNESCO. (2021, November 23). *Recommendation on the Ethics of Artificial Intelligence.* Retrieved April 1, 2022, from https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000381137&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_e86c4b5d-5af9-4e15-be60-82f1a09956fd%3F_%3D381137eng.pdf&locale=en&multi=true&ark=/ark:/48223/p

United Nations Human Rights Office of the High Commissioner. (2022, February 16). Better data collection bolsters human rights of marginalised people. Retrieved February 16, 2022, from https://www.ohchr.org/EN/NewsEvents/Pages/Data-and-human-rights.aspx

United Nations News. (2021, July 19). Pegasus: Human rights-compliant laws needed to regulate spyware. Retrieved January 3, 2022, from https://news.un.org/en/story/2021/07/1096142

Walters, W. P., & Murcko, M. (2020, January 30). Assessing the impact of generative AI on medicinal chemistry. *Nature Biotechnology , 38*, pp. 143-145. Retrieved February 22, 2022, from https://www.nature.com/articles/s41587-020-0418-2

Wilson, H. J. (2015, April 15). What Is a Robot, Anyway? *Harvard Business Review.* Retrieved February 16, 2022, from https://hbr.org/2015/04/what-is-a-robot-anyway

World Commission in the Ethics of Scientific Knowledge and Technology. (2019, February 26). *Preliminary Study on the Ethics of Artificial*

Intelligence. Retrieved April 2, 2022, from https://unesdoc.unesco.org/ark:/48223/pf0000367823

Yeung, K., Howes, A., & Pogrebna, G. (2020). AI Governance by human Rights-Centered Design, Deliberation and Oversight: An End to Ethics Washing. In M. d. Dubber, F. Pasquale, & S. Das, *The Oxford Handbook of Ethics of AI* (pp. 76-106). Oxford: Oxford University Press

NOTAS

1   The term was introduced in the 1950's and was referred to machines that could do more than routine tasks (World Commission in the Ethics of Scientific Knowledge and Technology. (2019, February 26). Preliminary Study on the Ethics of Artificial Intelligence. Retrieved April 2, 2022, from https://unesdoc.unesco.org/ark:/48223/pf0000367823, p. 3.)

2   The International Federation on Robotics distinguishes between industrial robots and service robots. The first is defined as «[…] an automatically controlled, reprogrammable multipurpose manipulator programmable in three or more axes.» The second is a robot that «[…] performs useful tasks for humans or equipment excluding industrial automation application.» (International Federation on Robotics. (n.d.). Topics and Definitions. Retrieved February 16, 2022, from https://ifr.org/#topics).

    Although, terminology and definition are not unanimous within the scientific community.

    Despite the no consensus, a lot has advanced since a sole master/slave definition. This notion evolved in 1920 by Karel Čapek's play, *Rossum's Universal Robots, and* «[…] *Čapek's robots were mass produced workers assembled from artificially synthesized organic material.*» (Wilson, H. J. (2015, April 15). What Is a Robot, Anyway? *Harvard Business Review*. Retrieved February 16, 2022, from https://hbr.org/2015/04/what-is-a-robot-anyway)

3   The study commissioned by the European Parliament's Legal Affairs Committee, European Civil Law Rules in Robotics, recalls for a common European definition of the various categories of autonomous robot. The notion of smart robot seems even harder. (Policy Department for "Citizens' Rights and Constitutional Affairs. (2016, October). European Civil Law Rules in Robotics. Retrieved February 17, 2022, from http://www.europarl.europa.eu/committees/fr/supporting-analyses-search.html, pp. 8-12).

4   The European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) raised important questions and, mostly, deep concerns surrounding robots and their behavior.

5   The terms of an AI policy and framework are also crucial to the functionality of robots.

6   Robots are increasingly assuming some important domains, For instance, according to the Science and Technology Directorate, can «[…] assist with enhancing capabilities of CBP [U.S. Customs and Border Protection] personnel, while simultaneously increasing their safety downrange.» (Science and Technology Directorate. (2022, February 1). Feature Article: Robot Dogs Take Another Step Towards Deployment at the Border. Retrieved February 20, 2022, from https://www.dhs.gov/science-and-technology/news/2022/02/01/feature-article-robot-dogs-take-another-step-towards-deployment).

7   Traditionally medicine is reactive. AI is enhancing P4 Medicine. A predictive, preventive, personalized and participatory medicine. Pioneers of this approach to health care are Leroy Hood, James Heath, Michael Phelps, and Biaoyang Lin. In 2004, they have published a paper on Science about technologies enable predictive and Preventative Medicine. It will lead to a personalized medicine. See Hood, L., Heath, J. R., Phelps, M. E., & Lin, B. (2004, October 22). Systems Biology and New Technologies Enable Predictive and Preventative Medicine. Science, 306, Issue 5696, pp. 640-643. DOI: 10.1126/science.1104635

    In 2021, researchers from Kaunas University of Technology and Vytautas Magnus University in Lithuania developed a deep learning-based method that can predict mild cognitive impairment and Alzheimer's disease with an accuracy of over 99%. See Odusami, M., Maskeliūnas, R., Damaševičius, R., & Krilavičius, T. (2021, June 10). Analysis of Features of Alzheimer's Disease: Detection of Early Stage from Functional Brain

Changes in Magnetic Resonance Images Using a Finetuned ResNet18 Network. *Diagnostics, 11 (6)*. https://doi.org/10.3390/diagnostics11061071.

[8] Although AI in biotechnology industry does not relate solely to programs that predict outcomes. We can also be used in drug discovery and development. Walters and Murcko published an article on Nature Biotechnology assessing the impact of generative AI on medicinal chemistry. See Walters, W. P., & Murcko, M. (2020, January 30). Assessing the impact of generative AI on medicinal chemistry. Nature Biotechnology , 38, pp. 143-145. Retrieved February 22, 2022, from https://www.nature.com/articles/s41587-020-0418-2.

[9] Neuroscientists are also testing the limits creating a symbiotic relation between AI and neuroscience. A clear example are the research developments of AI to enable paralyzed people to use robotic limbs as their own. Researcher like Chethan Pandarinath are recording brain activity from paralyzed people hopping to «[…] identify the patterns of electrical activity in neurons that correspond to a person's attempt to move their arm in a particular way, so that the instructions can then be fed to a prosthesis.» (Savage, N. (2019, August 21). How AI and neuroscience drive each other forwards. Nature. Retrieved February 22, 2022, from https://www.nature.com/articles/d41586-019-02212-4).

[10] Also, robotic chirurgic.

[11] For the information society, in a related topic, the internet of things (IoT). The IoT is a «[…] global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.» (International Telecommunication Union. (2012, June 15). Overview of the Internet of things. Recommendation ITU-T Y.2060. Retrieved March 23, 2022, from https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=y.2060, p. 1).

[12] «[R]efers to large amount of data produced very quickly by a number of diverse sources. Data can either be created by people or generated by machines, such as sensors gathering climate information, satellite imagery, digital pictures and videos, purchase transaction records, GPS signals, and more.» (European Commission. (2021, December 13). Big Data. Retrieved February 15, 2022, from https://digital-strategy.ec.europa.eu/en/policies/big-data).

[13] In the article published by Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma e Sandeep Kaushi the «International Data Corporation (IDC) estimated the approximate size of the digital universe in 2005 to be 130 exabytes (EB). The digital universe in 2017 expanded to about 16,000 EB or 16 zettabytes (ZB). IDC predicted that the digital universe would expand to 40,000 EB by the year 2020» (Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. Journal of Big Data, 6:54, 1-25. doi: https://doi.org/10.1186/s40537-019-0217-0, p. 3).

[14] European Commission, Communication from the Commission on Artificial Intelligence for Europe COM(2018) 237 final, 2018.

[15] Automated machine learning «[…] refers to the tools and processes which make it easy to build, train, deploy and serve custom machine learning models.» (Qlik. (n.d.). AutoML. What it is, why you need it, and best practices. Retrieved February 20, 2022, from https://www.qlik.com/us/augmented-analytics/automl).

[16] The «[…] use of algorithms that find a pattern in data without explicit instructions» is machine-learning and «[a] system might learn how to associate features of inputs such as images with outputs such as labels.» (Hutson, M. (2017, July 7). AI Glossary: Artificial intelligence, in so many words. Science, 357, Issue 6346, https://www.science.org/doi/10.1126/science.357.6346.19, p. 19).

The algorithms are a «*finite sequence of formal rules (logical operations and instructions) making it possible to obtain a result from the initial input of information. This sequence may be part of an automated execution process and draw on models designed through machine learning.*» (European Commission for the Efficiency of Justice. (2018, December 3-4). *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment.* Retrieved April 3, 2022, from https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c, p. 69).

Machine learning

«[M]*akes it possible to construct a mathematical model from data, incorporating a large number of variables that are not known in advance. The parameters are configured gradually during the learning phase, which uses training*

data sets to find and classify links. The different methods of machine learning are chosen by the designers depending on the nature of the tasks to be completed (grouping). These methods are usually classified into three categories: (human) supervised learning, unsupervised learning and reinforcement learning. These three categories group together different methods including neural networks, deep learning, etc.» (Ibid, p. 72).

ML can be sub-categorized into three types of learning: supervised learning, unsupervised learning, and reinforced learning. (Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An Introduction to Ethics in Robotics and AI.* Springer, p. 25).

[17] We can also include some subfields as

«[…] *deep learning, which uses neural networks to identify complex patterns in high-volume data, cognitive computing, which is used to simulate the functioning of the human brain to solve complex problems, and natural language processing, which helps computers understand and interpret human language.*» (Qlik. (n.d.). *Big Data AI. How Big Data and AI Work Together.* Retrieved February 20, 2022, from https://www.qlik.com/us/augmented-analytics/big-data-ai).

[18] European Comission's High-Level Expert Group on Artificial Intelligence. (2018, December 18). Retrieved February 14, 2022, from https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf, p. 1

[19] Annex I, referred in Article 3 , includes:

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c) Statistical approaches, Bayesian estimation, search and optimization methods.

European Commission, Proposal for a Regulation on the European Parliament and the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial intelligence Act) and Amending Certain Union Legislative Acts COM(2021) 206 final, 2021.

[20] Ibid.

[21] Opinion of the European Committee of the Regions. (2022, February 28). European approach to artificial intelligence - Artificial Intelligence Act (revised opinion) (2022/C97/12). Retrieved March 12, 2022, from Official Journal of the European Union (C97/602).

[22] Ibid.

[23] Ibid.

[24] For instance, the USA's National Artificial Intelligence Initiative Act of 2020 defines AI as

«[…] *a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to—*

(A) perceive real and virtual environments;

(B) abstract such perceptions into models through analysis in an automated manner; and

(C) use model inference to formulate options for information or action.» (National Artificial Intelligence Initiative Act of 2020. (2022, March 3). Retrieved from https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210, Division E, Section 5002, (3))

[25] In the 1950's Alan Turing suggested that machine intelligence is based upon its ability to exhibit intelligence which is undistinguishable from the and intelligent human's behavior. (Bartneck, Lütge, Wagner, & Welsh, op. cit. p. 23).

[26] The communication from the European Commission on 2018 reasons a AI strategy for Europe making «trust» a prerequisite to ensure the referred «human-centric approach to AI» European Commission, Building Trust in Human-Centric Artificial Intelligence (COM(2019) 168 final, 2019.

[27] Also referred in the Briefing of the European Parliament on the Artificial Intelligence Act (AI Act). European Parliament, Artificial intelligence Act, 2021.

[28] The OECD Council Recommendation on Artificial Intelligence underlines the principle of «human-centred values and fairness» for responsible stewardship of trustworthy AI. (OECD Council. (2019, May 22).

Recommendation on Artificial Intelligence. Retrieved January 14, 2022, from https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449).

29 See the draft of the AI Act.

30 National Artificial Intelligence Initiative Act of 2020, op. cit.,. Division E, Title LI, Section 5101, (a)

31 The document can be retrieved from https://www.congress.gov/bill/117th-congress/senate-bill/3572/text?r=1&s=1 [last accessed April 12, 2022]

32 The document can be retrieved from https://www.congress.gov/bill/117th-congress/house-bill/6580/text?r=37&s=1 [last accessed April 12, 2022].

33 Morgan, C. S., Langlois, F., & Lan, J. (2022, April 8). U.S. House and Senate Reintroduce the Algorithmic Accountability Act Intended to Regulate AI. Retrieved April 12, 2022, from Lexology: https://www.lexology.com/library/detail.aspx?g=c9bfbfca-a88e-4fd9-b556-d903a4cf0bd7

34 Declaration of cooperation on Artificial Intelligence (AI), 2018.

35 European Commission. (2018, April 25). Communication from the Commission on Artificial Intelligence for Europe COM(2018) 237 final. Retrieved February 14, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN#footnote12. Point 3.1. of the communication «Boosting the EU's technological and industrial capacity and AI uptake across economy».

36 The Commission declared it would, in the period of 2018-20, around EUR 1.5 billion and aimed to stimulate more private investments in AI under the European Fund for Strategic Investments and, for the period 2021-2027, the Commission proposals will open the door to:

«· *upgrading the pan-European network of AI excellence centres;*

· *research and innovation in fields such as explainable AI , unsupervised machine learning, energy and data efficiency;*

· *additional Digital Innovation Hubs, world-leading testing and experimentation facilities in areas such as transport, healthcare, agrifood and manufacturing, supported by regulatory sandboxes;*

· *supporting the adoption of AI by organisations across all sectors, including public interest applications, through co-investment with Member States;*

· *exploring joint innovation procurement for the use and development of AI; and*

· *a support centre for data sharing, which will be closely linked with the AI-on-demand platform to facilitate development of business and public sector applications.*» (European Commission, Communication from the Commission on Artificial Intelligence for Europe COM(2018) 237 final, op. cit. Point 3.1. of the communication «Boosting the EU's technological and industrial capacity and AI uptake across economy»).

37 For the period of 2019-2024.

38 Leyen, U. v. (2019). Political Guidelines for the next European Commission 2019-2024. Retrieved March 9, 2022, from https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_en_0.pdf, p. 13.

39 EU-US Trade and Technology Council Inaugural Joint Statement. (2021, September 29). Retrieved March 8, 2022, from https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_21_4951.

40 European Commission. (2020, February 19). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committe of the Regions Shaping Europe's digital future (COM/2020/67 final). Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0067.

41 European Commission. (2020, February 19). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions A European strategy for data (COM/2020/66 final). Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066.

42 European Commission. (2019, April 8). Building Trust in Human-Centric Artificial Intelligence (COM(2019) 168 final). Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52019DC0168&qid=1620297270190.

43 European High Performance Computing Joint Undertaking. (n.d.). Retrieved March 26, 2022, from https://eurohpc-ju.europa.eu/discover-eurohpc-ju.

44 European Commission. (2020, February 19). White Paper on Artificial Intelligence - a European approach to excellence and trust COM(2020) 65 final. Retrieved February 15, 2022, from https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

45 Ibid, p. 25.

46 The European Commission established the AI Alliance forum to enhance a broad and public debate around AI matters. The forum can be accessed at https://futurium.ec.europa.eu/en/european-ai-alliance [last accessed April 19, 2022].

47 European Commission. (2020, February 19). Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics COM/2020/64 final. Retrieved February 15, 2022, from https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1593079180383&uri=CELEX%3A52020DC0064, pp. 16-17.

48 European Commission. (2021, June 30). Adapting liability rules to the digital age and circular economy. Retrieved March 25, 2022, from file:///C:/Users/almeidaaraujo/Downloads/090166e5df292bd3.pdf.

49 European Commission. (2021, April 21). Communication from the Commission to the European Parliament, the council, the European Ecnomic and Social Committee and the Committee of Regions: Fostering a European approach to Artificial Intelligence (COM(2021) 205 final). Retrieved April 2, 2022, from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM%3A2021%3A205%3AFIN .

50 European Parliament. (2020, October 20). Resolution on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html.

51 European Parliament. (2020, October 20). Resolution on a civil liability regime for artificial intelligence, (2020/2014(INL)). Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html.

52 European Parliament. (2020, October 20). Resolution on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)). Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_EN.html.

53 European Parliament. (2021, October 6). Artificial intelligence in criminal law and its use by the police and judicial (2020/2016(INI)). Retrieved January 5, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.pdf.

54 European Parliament. (2021, May 19). Resolution on artificial intelligence in education, culture and the audio-visual sector (2020/2017(INI)). Retrieved January 7, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_EN.html .

55 European Parliament. (2020, June 18). Decision of 18 June 2020 on setting up a special committee on artificial intelligence in a digital age, and defining its responsibilities, numerical strength and term of office (2020/2684(RSO)). Retrieved January 6, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2020-0162_EN.html.

56 Independent High-Level Expert Group on Artificial Intelligence to the European Commission. (2019, April 8). Ethics Guidelines for a Trustworthy AI. Brussels. Retrieved February 15, 2022, from https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf.

57 Independent High-Level Expert Group on Artificial Intelligence to the European Commission, 2019, ibid, p. 2.

58 Independent High-Level Expert Group on Artificial Intelligence to the European Commission, ibid, followed by Policy and investment recommendations for trustworthy AI.

59 Proposal for a Regulation on the European Parliament and the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial intelligence Act) and Amending Certain Union Legislative Acts COM(2021) 206 final.

60 European Commission. (2021, April 21). Proposal for a Regulation on the European Parliament and the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial intelligence Act) and Amending Certain Union Legislative Acts COM(2021) 206 final. Retrieved December 18, 2021, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.

61 Three levels of risk: unacceptable risk; high risk; and low/minimal risk.

Under Title II of the AI Act, article 5 lists prohibited AI systems considering the unacceptable risk to society, including systems that distort a person's behavior, exploits any vulnerabilities of a specific group, social scoring.

High-Risk AI Systems are classified in article 6 and are considered as such when fulfilling two conditions: *«(a) the AI system is intended to be used as a safety component of a product, or is itself a product, covered by the Union harmonisation legislation listed in Annex II;*

(b) the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation legislation listed in Annex II.»

These AI systems are not prohibited although a strict regulation is proposed.

The low/minimal risk AI systems are the ones that are not prohibited nor the ones that are considered high-risk. Although the Act encourages, in its article 69, the

«[…] drawing up of codes of conduct intended to foster the voluntary application to AI systems other than high-risk AI systems of the requirements set out in Title III, Chapter 2 on the basis of technical specifications and solutions that are appropriate means of ensuring compliance with such requirements in light of the intended purpose of the systems»

62 Included at the Charter of Fundamental Rights of the European Union that brings together personal, civic, political, and also social and economic rights.

63 Set out in Article 2 of the Treaty on European Union.

64 European Commission, White Paper on Artificial Intelligence, p. 25.

65 Shared by the US and included in the joint statement EU-US TTC that covered, among others, the issue of «Trade Policy Cooperation towards Non-Market Economies (NMEs). EU-US Trade and Technology Council Inaugural Joint Statement, 2021.

66 Powers, T. M., & Ganascia, J.-G. (2020). The Ethics of the Ethics AI. In M. D. Dubber, F. Pasquale, & S. Das, The Oxford Handbook of Ethics of AI (pp. 26-51). Oxford: Oxford University Press, p. 29.

67 Müller-Abdelrazeq, S. L., Schönefeld, K., Haberstroh, M., & Hees, F. (2019). Interacting with Collaborative Robots—A Study on Attitudes and Acceptance in Industrial Contexts. In O. Korn, Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction (pp. 101-117). Springer.

68 The European Commission makes it explicit about the immaterial harms that AI humanoid robots can cause, and not only physical harm, but also mental health issues for users, in particular vulnerable users (such as elderly persons) for instance, excessive stress and discomfort. European Commission, Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics COM/2020/64 final, 2020, p. 8.

69 Recalling that genetics, nanotechnology, and robotics are, for the author, the most powerful 21st century technologies. (Joy, B. (2000, April 1). Why the Future Doesn't Need Us. Retrieved April 20, 2022, from Wired: https://www.wired.com/2000/04/joy-2/)

70 Ibid.

71 Jonas, H. (1984). The Imperative of Responsability: In Search of an Ethics for the Technological Age. Chicago: The University of Chicago Press, p. 9.

72 Sarah L. Müller-Abdelrazeq, Kathrin Schönefeld, Max Haberstroh and Frank Hees, op. cit., analyse the quality of interaction and attitudes especially with collaborative robots.

73 A particular interesting study about this topic is presented by Sirithunge, H. P., Bandara, H. M., Jayasekara, A. G., Chandima, D. P., & Abeykoon, H. M. (2019). A Study on Robot-Initiated Interaction: Toward Virtual Social Behavior. In O. Korn, Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction (pp. 33-70). Springer.

74 Barrat, J. (2013). Our Final Invention: Artificial intelligence and the end of the human Era. New York: St. Martin's Press, p. 30.

75 As quoted by Monlezun, D. J. (2020). The Global Bioethics of Artificial Intelligence and Human Rights. Cambridge: Cambridge Schoolar Publishing, p. 282, quoting James Barrat, 2013, op. cit.

76 Monlezun, op. cit., p. 282

77 Bostrom, N. (2017 (first published in 2014)). Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford Unversity Press, p. 53.

78 World Commission in the Ethics of Scientific Knowledge and Technology, Preliminary Study on the Ethics of Artificial Intelligence, op. cit., p. 3.

79 The neurons or neural networks «[…] are computing systems vaguely inspired by the biological neural networks that constitute animal brains.» (European Commission for the Efficiency of Justice, European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, 2018, p. 73).

80 Deep learning is based on ANNs of several layers.
    The layers «[…] *enables the machine to recognize complex concepts such as human faces, human bodies, speech understanding and all types of images classification.*» (World Commission in the Ethics of Scientific Knowledge and Technology, Preliminary Study on the Ethics of Artificial Intelligence, op. cit, p. 7).

81 European Commission for the Efficiency of Justice, European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, 2018, p. 73.

82 In the 1980's John Searle divided AI in two categories. Weak and strong AI, depending on if AI is limited to a single task or programmed computer that would have a mind in exactly the same sense of a human mind, something like a general intelligence. (Bartneck, Lütge, Wagner, & Welsh, op. cit., p. 24).

83 Ibid, p. 44.

84 Considering the subject, we can divide AI two categories: narrow AI (ANI) and general AI (AGI).
    «*An AGI, like us, could drive a car, have a conversation, play chess, and read and produce a summary of a paper. Perhaps not all at the same time, but importantly AGI is not limited to a single domain. We currently do not have an AGI, and A.I. researchers do not have clarity regarding when or if we will. […] Narrow artificial intelligence is often talked about as deep learning and sometimes as machine learning* […]. » (Reed, R. (2021, May 31). A.I. in Religion, A.I. for Religion, A.I. and Religion: Towards a Theory of Religious Studies and Artificial Intelligence. *Religions, 12*. doi: https://doi.org/10.3390/rel12060401, p. 3).
    ANI performs four different tasks: «classification, prediction, generation (Reed quotes Mitchell, Melanie. 2020. Artificial Intelligence: A Guide for Thinking Humans. Stuttgart: Picador) and also «[…] *reinforcement learning, which attempts to learn through trial and error by maximizing a reward signal*» Ibid.

85 The European Commission Joint Research Centre addresses the AI dichotomy. I.e. ANI and AGI. Joint Research Centre. (2019). 1st Workshop on Artificial Intelligence at JRC, Ispra 23rd May 2018. doi:10.2760/705074, p. 68.

86 Bartneck, Lütge, Wagner, & Welsh, op. cit., p. 42.

87 See Gellers, J. C. (2021). Rights for Robots: Artificial Intelligence, Animal and Environmental Law. New York: Routledge.

88 Agamben, G. (1998 (originally published in 1995)). Homo Sacer: Sovereign Power and Bare Life (2nd ed.). (D. Heller-Roazen, Trans.) USA: Stanford University Press.

89 Liao, S. M. (2020). The Moral Status and Rights of Artificial Intelligence. In S. M. Liao, Ethics of Artificial Intelligence (pp. 480-503). Oxford: Oxford University Press, p. 489.

90 Monlezun, op. cit., 2020.

91 Eduardo Figueiredo states that the Genetic, Nanotechnology and Robotics (GNR) is a transformation program of the human body and, at the very limit, it can include a passage to scenarios of cyborgs. (Figueiredo, E. A. (2021). Direito e Nanobiotecnociência. Reflexões na encruzilhada da inovação, do risco e da crise do(s) direito(s). Coimbra: Almedina, p. 29).

92 Ormond, K. E., Mortlock, D. P., Scholes, D. T., Bombard, Y., Brody, L. C., Faucett, W. A., . . . Young, C. E. (2017, August 3). Human Germline Genome Editing. The American Journal of Human Genetics, 101 (2), pp. 167-176. https://doi.org/10.1016/j.ajhg.2017.06.012, p. 168.

93 Fukuyama, F. (2002). Our Posthuman Future. Consequences of the biotechnology revolution. New York, NY: Picador, and Buchanan, A. (2011). Beyond humanity? Oxford: Oxford University Press

94 Silver, M. L. (2007 (First Published 1997)). Remaking Eden: How Genetic Engineering and Cloning will Transform The American Family. USA: Ecco.

95 Fukuyama, op. cit., p. 94.

96 Agar, N. (2014). Truly Human Enhancement. A Philosophical Defense of Limits. Massachusetts: MIT Press, p. 189.

97 The first transhumanist declaration of Fereidoun Espandiary from 1978 states that «Nous voulons accélérer l'avancée de l'humanité jusqu'à la prochaine étape de son évolution» (Terence, M. (2016). Le Transhumanisme est un Intégrisme. Paris: Les Éditions du Cerf, p. 28).

98 Sève, L. (2006). Qu'est-ce que la personne humaine? Paris: La Dispute, p. 81.

99 Habermas, J. (2008). The Future of Human Nature. Cambridge: Polity Press, p. 28.

100 Bostrom, N. (2003). The Transhumanist FAQ - A General Introduction - Version 2.1 . Retrieved September 9, 2020, from World Transhumanist Association: https://www.nickbostrom.com/views/transhumanist.pdf.

101 Singularity «[…] is a religion created by people who have the experience of using computation to solve problems heretofore considered impossibly complex for machines.» (Ito, J. (2018, December 2). Resisting Reduction: A Manifesto. Journal of Design and Science. doi:10.21428/8f7503e4, p. 4).

102 Although not immortal, death is not necessarily the result of aging.

103 Ito, op. cit., p. 4.

104 Bostrom, 2017, op. cit., pp. 53-54.

105 One of the clearest examples was the national eugenic programme implemented under the Nazi racial hygiene policies.

106 More information about the project can be found in https://www.nextrembrandt.com/ [last accessed April 14, 2022]

107 More about the unfinished Symphonie Project can be found in https://consumer.huawei.com/au/campaign/unfinishedsymphony/ [last accessed April 14, 2022]

108 Immediately one question pops. Who is the author?

109 International Committee of the Red Cross. (2021, May 12). ICRC position on autonomous weapon systems. Retrieved May 3, 2022, from https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems.

110 Ibid.

111 Among others see Asaro, P. (2020). Autonomous Weapons and the Ethics of Artificial Intelligence. In S. M. Liao, , Ethics of artificial Intelligence (pp. 212-236). Oxford: Oxford University Press, p. 216.

112 Kumm, M. (2013). The Cosmopolitan Turn in Constitutionalism: An Integrated Conception of Public Law. Indiana Journal of Global Legal Studies, 20:2, 605-628. Retrieved April 14, 2022, from https://www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=1516&context=ijgls, p. 607.

113 Nemitz, P. (2018, August 14). Constitutional democracy andtechnology in the age of artificial intelligence. Phil. Trans. R. doi:http://dx.doi.org/10.1098/rsta.2018.0089.

114 Google, Facebook, Microsoft, Apple, and Amazon.

115 Nemitz, op. cit., p. 3.

116 Yeung, K., Howes, A., & Pogrebna, G. (2020). AI Governance by human Rights-Centered Design, Deliberation and Oversight: An End to Ethics Washing. In M. d. Dubber, F. Pasquale, & S. Das, The Oxford Handbook of Ethics of AI (pp. 76-106). Oxford: Oxford University Press, p. 78.

117 AI has already a profound influence in many aspects of our lives and concerns related to the lack of transparency and/ or information, discrimination, or bias, affect people's trust on AI systems.

118 Kalpokas, I., Kalpokiene, J. (2022). Deepfakes: A Realistic Assessment of Potentials, Risks, and Policy Regulation, Switzerland: Springer, p. 55.

119 Nadler, A., Crain, M. & Donavan, J. (2018). Weaponizing the digital influence machine: The Political Perils of Online Ad Tech. Data & Society Research Institute, https://datasociety.net/library/weaponizing-the-digital-influence-machine//.

120 Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. Journal of Economic Perspectives, 31 (2), 211-236. doi=10.1257/jep.31.2.211, p. 212.

121 The authors data base «[…] contains 115 pro-Trump fake stories that were shared on Facebook a total of 30 million times, and 41 pro-Clinton fake stories shared a total of 7.6 million times.» Allcott & Gentzkow, op. cit., p. 212.

122 Allcott & Gentzkow, op. cit., p. 232.

123 See https://www.reuters.com/article/uk-factcheck-nancypelosi-manipulated-idUSKCN24Z2BI [last accessed May 2, 2022]

124 See https://nypost.com/2022/03/17/deepfake-video-shows-volodymyr-zelensky-telling-ukrainians-to-surrender/ [last accessed May 6, 2022]

125 Franke, U. (2021, June). Artificial Intelligence diplomacy: Artificial Intelligencegovernance as a new European Union external policy tool. Retrieved April 28, 2022, from Study Requested by the AIDA committee: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662926/IPOL_STU(2021)662926_EN.pdf.

126 Annex to the European Commission. (2021). Coordinated Plan on Artificial Intelligence 2021 Review. Retrieved April 2, 2022, from https://digital-strategy.ec.europa.eu/pt/node/9757.

127 For instance, the AI Watch. The Commission's Joint Centre works with the Member States and just released AI Watch Index that provides an analysis of indicators related to AI «[…] (i) global view on the AI landscape, (ii) industry, (iii) research and development (R&D), (iv) technology, and (v) societal aspects.» (Joint Research Centre; European Commission. (2022, March 28). AI Watch Index 2021. Retrieved April 18, 2022, from https://ai-watch.ec.europa.eu/publications/ai-watch-index-2021_en, p. 3).

128 European Commission, Coordinated Plan on Artificial Intelligence 2021 Review, 2021, p. 2.

129 Truing and his team constructed the first operational computer named Robinson. The group relevant important help to the British in World War II making transcriptions of Nazi messages. Kurzweil, R. (2000). The Age of Spiritual Machines: When Computers Exceed Human Intelligence . New York: Penguin books, p. 115.

130 Turing, A. (1950, October). Computing Machinery and Intelligence. Mind. A Quarterly Review of Psycology and Philosophy, LIX (236), 433-460. Retrieved April 20, 2022, from https://watermark.silverchair.com/lix-236-433.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAu YwggLiBgkqhkiG9w0BBwagggLTMIICzwIBADCCAsgGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQ QMvKEkFW4xrg57Pl5CAgEQgIICmdlXdLEuv-7V6ssL2ndgapYY5NCh343GhNtuWUDTkhx.

131 The new form of the problem is described by Truing as an «imitation game», which it is played

  «[…] with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'.» Ibid, p. 433.

132 Ibid, p. 434.

133 Ibid, p. 456.

134 Turing recognized this both direction communication. Ibid, p. 456.

135 Kurzweil, op. cit., p. 157.

136 Ibid.

137 Saygin, A. P., Cicekli, I., & Akman, V. (2000, October). Turing Test: 50 Years Later. Minds and Machines, 10(4), 463-518. doi:10.1023/A:1011288000451, p. 512.

138 Kurzweil, op. cit., p. 159.

139 Ibid, pp. 159-60.

140 Ammanath, B. (2022). Trustworthy AI: A Business Guide for Navigating Trust and Ethics in AI. New Jersey: Wiley, p. 47.

141 James Barrat explains how he thinks the law contain echoes, besides from the Hippocratic oath, of the Judeo-Christian rule «Thou Shalt Not Kill» and the «[…] notion of that sin results from acts committed and omitted […]» (Barrat, J. (2013). Our Final Invention: Artificial intelligence and the end of the human Era. New York: St. Martin's Press, p. 20).

142 Policy Department for "Citizens' Rights and Constitutional Affairs, 2016, p. 13.

143 In general, we can talk about physical harm, but also, economic, environmental, and legal harm. About this topic see Ammanath, op. cit.

144 Greene, D., Hoffmann, A. L., & Stark, L. (2019). Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. Proceedings of the 52nd Hawaii International Conference on System Sciences, 2122-2131. Retrieved April 13, 2022, from https://hdl.handle.net/10125/59651, p. 2124.

145 Michelle Bachelet was speaking on 14 September 2021 at a Council of Europe hearing on the implications stemming controversy over Pegasus spyware. Bachelet, M. (2021, September 14). Committee on Legal Affairs and Human Rights, Parliamentary Assembly Council of Europe. Retrieved January 3, 2022, from https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=27455&LangID=E.

146 United Nations News. (2021, July 19). Pegasus: Human rights-compliant laws needed to regulate spyware. Retrieved January 3, 2022, from https://news.un.org/en/story/2021/07/1096142.

147 OCCRP. (2021). The Pegasus Project. Retrieved May 15, 2022, from https://www.occrp.org/en/the-pegasus-project/

148 Bachelet, op. cit.

149 Marczak, B. B., Scott-Railton, J., Berdan, K., Razzak, B. A., & Deibert, R. (2022, January 3). Hooking Candiru: Another Mercenary Spyware Vendor Comes into Focus. Retrieved July 15, 2021, from https://citizenlab.ca/2021/07/hooking-candiru-another-mercenary-spyware-vendor-comes-into-focus/.

150 Jagland, T. (2019, May 16-17). Ready for future challenges – reinforcing the Council of Europe. Helsinki. Retrieved February 14, 2020, from https://rm.coe.int/168093af03, p. 9.

151 A form of predictive justice which is «[...] the analysis of large amounts of judicial decisions by artificial intelligence technologies in order to make predictions for the outcome of certain types of specialised disputes.» (European Commission for the Efficiency of Justice, 2018, European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, p. 74).

152 See Almeida Araújo, M., & Meireis, A. (2020). A Century After Plessy v. Ferguson. The Struggle for Racial Equality in the Algorithm Era. Lusiada Law Review, 23/24, pp. 195-204.

153 Larson, Mattu, Kirchner, & Angwin, 2016, How We Analyzed the COMPAS Recidivism Algorithm.

154 Larson, Mattu, Kirchner, & Angwin, op. cit.

155 Some might be embedded, even unconscientiously, in our societies

156 In 2019 it was adopted by the Committee of Ministers at the 1337th meeting of the Ministers' Deputies the Declaration on the manipulative capabilities of algorithmic processes. The Committee to the growing threat to «[…] *the right of human beings to form opinions and take decisions independently of automated systems, which emanates from advanced digital technologies. Attention should be paid particularly to their capacity to use personal and non-personal data to sort and micro-target people, to identify individual vulnerabilities and exploit accurate predictive knowledge, and to reconfigure social environments in order to meet specific goals and vested interests*;» (Committee of Ministers. (2019, February 13). Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes. *Decl(13/02/2019)1*. Retrieved February 16, 2022, from https://search.coe.int/cm/pages/result_details.aspx?ObjectId=090000168092dd4b).

157 The data diet its particularly important. For instance, in March 2016, Microsoft launched Tay, a chatbot meant to improve Microsoft's understanding of conversation language. Although, in 24 hours the technology incorporated prejudices and started tweeting racist, misogynist, and other discriminatory remarks.

158 The belief that numbers cannot lie and always represent objective truth. (Executive Office of the President. (2016, May). Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights. Retrieved March 25, 2022, from https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf, p. 10).

159 State of Wisconsin v. Loomis. Retrieved March 21, 2022, from https://casetext.com/pdf-email?slug=state-v-loomis-22

160 Freeman, K. (2016, 12). Algorithmic Injustice: How The Wisconsin Supreme Court Failed To Protect Due Process Rights in State v. Loomis. North Carolina Journal of Law & Technology, 18, pp. 75-106, pp. 89-89.

161 Adopted at the 31st plenary meeting of the CEPEJ (Strasbourg, 3-4 December 2018)
European Commission for the Efficiency of Justice, European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, 2018.

162 In Europe, Harm Assessment Risk Tool (HART) was developed in partnership with Cambridge University. This technology based on machine learning was trained using Durham Police archives dating from 2008 to 2012. European Commission for the Efficiency of Justice, 2018, European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, p. 51.

163 Ibid, p. 55.

164 European Parliament. (2021, October 6). Resolution on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)). Retrieved April 15, 2022, from https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html

165 Ibid.https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html

166 It seems hard to understand, for example, granting humanoid robot Sophia the Saudi Arabian citizenship in 2017 and women struggle against discrimination. Not much more than a publicity stunt Sophia represents, however, the mystique around robots.

   In a very interesting perspective Joshua C. Gellers presents a theory about the *status* robots and question «*can robots have rights*?» (Gellers, J. C. (2021). *Rights for Robots: Artificial Intelligence, Animal and Environmental Law.* New York: Routledge, p. 15).

   In fact, we can understand the responsibilities that we have to sentient beings and «[a]*ll currently existing artificial entities are nonsentient but – unlike a rock – their interactions and designs evoke the impression of conscious entities with personalities and emotions.*» (Donath, J. (2020). Ethical Issues in our Relationship with Artificial Entities. In M. D. Dubber, & S. D. Frank Pasquale, *The Oxford Handbook of Ethics of AI* (pp. 52-73). Oxford: Oxforxd University Press, p. 53).

   This topic is quite interesting, OECD and the UK have been recommending that «[…] *AI should never deceptively appear to be human.*» B(ryson, 2020, The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law and Regulation, p. 22). The quote is referring to the OECD AI ethics principles (OECD Council, 2019, *Recommendation on Artificial Intelligence*).

167 For instance, racial representation emerged by research «[…] demonstrating facial recognition technology's difficulty recognizing people of color (especially black people), companies like Microsoft began touting "inclusivity" efforts aimed at improving facial recognition's performance across skin tones.» (Greene, Hoffmann, & Stark, op. cit., p. 2123).

168 AI and algorithms are used in «decision making» tools and questions related to human rights led the Committee of Ministers of the Council of Europe to adopt a Declaration on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of social safety net on 17 March 2021 at the 1399th meeting of Ministers' Deputies.

   The Committee urged Member states to have special attention, amongst other recommendation, to «*the possible risks to human rights, including social rights, that might follow from the use of computer-assisted or AI-enabled decision making by public authorities in the area of social services, when no qualified supervision is guaranteed, in particular as regards the management, attribution or revocation of entitlements, assistance and related benefits;*». (Committee of Ministers. (2021, March 17). Declaration by the Committee of Ministers on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of the social safety net. *Decl(17/03/2021)2.* Retrieved February 16, 2022, from https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680a1cb98).

   It also should be ensured decision-making systems comply with the legal principles of certainty, legality, data quality, non-discrimination and transparency.

169 She explains her experience in the airport security protocol. (Costanza-Chock, S. (2018, July 18). Design Justice, A.I., and Escape from the Matrix of Domination. Journal of Design and Science. doi:10.21428/96c8d426).

170 Ibid, p. 4.

171 Gebru, T. (2020). Race and Gender: Data-Driven Claims about Race and Gender Perpetuated the Negative Biases of the Day. In M. D. dubber, F. Pasquale, & S. Das, The Oxford Handbook of ethics of AI (pp. 252-269). Oxford: Oxford University Press, p. 259.

172 Molnar, P. (2019, December). Technology on the margins: AI and global migration management from a human rights perspective. Cambridge International Law Journal, 305-330. doi:https://doi.org/10.4337/cilj.2019.02.07.

173 Crenshaw, K. W. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. Stanford Law Review, 43(6), 1241-1299.

174 Costanza-Chock, op. cit, p. 7.

175 The news can be found in Reuters webpage at https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G [last accessed April, 29, 2022].

176 Ibid.

177 Healthcare care services and pharmaceutical industry, but also in financial areas, synthetic data has been used considering that, with synthetic data, the issues data sharing, and data privacy.

178 To increase trust in synthetic data some considerations must be addressed: «The development of a privacy assurance report to remove concerns of re-identification. Utility measurements and assessment of the quality of synthetic data to enable generation.» (James, S., Harbron, C., Branson, J., & Sundler, M. (2021, December 13). Synthetic data use: exploring use cases to optimise data utility. Discover Artifcial Intelligence, 1(15). https://doi.org/10.1007/s44163-021-00016-y).

179 Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights from the Executive Office of the President, 2016.

180 Executive Office of the President, 2016, p. 77

181 Ibid, p. 9.

182 Committee of Ministers, 2019, Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes.

183 Arun, C. (2020). AI and the Global South: Designing for Other Worlds. In M. D. Dubber, F. Pasquale, & S. Das, The Oxford Handbook of Ethics of AI (pp. 588-607). Oxford: Oxford University Press, p. 591.

184 Ammanath, op. cit., p. 44, quoting the definition of «AI robustness» of The International Organization for Standardization.

185 United Nations Human Rights Office of the High Commissioner. (2022, February 16). Better data collection bolsters human rights of marginalised people. Retrieved February 16, 2022, from https://www.ohchr.org/EN/NewsEvents/Pages/Data-and-human-rights.aspx.

186 Ad Hoc Committee on Artificial Intelligence. (2020, September 23). Progress report. Retrieved March 17, 2022, from https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016809ed062.

187 The CAHAI has fulfilled its mandate and has been succeeded by the Committee on Artificial Intelligence (CAI)

188 Ben-Israel, I., Cerdio, J., Ema, A., Friedman, L., Ienca, M., Mantelero, A., . . . Vayena, E. (2020, December). Towards a 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. Retrieved April 10, 2022, from Compilation of contributions prepared by the CAHAI Secretariat: https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a.https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a

189 Ad Hoc Committee on Artificial Intelligence. (2020, December 17). Feasibility Study. Retrieved April 14, 2022, from https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da.https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da

190 Leslie, D., Burr, C., Aitken, M., Cowls, J., Katell, M., & Briggs, M. (2021). Artificial Intelligence, human rights, Democracy, and the rule of Law: a primer. The Council of Europe. Retrieved April 3, 2022, from https://rm.coe.int/primer-en-new-cover-pages-coe-english-compressed-2754-7186-0228-v-1/1680a2fd4a.

191 Ibid.

192 There are several examples. For instance, Google (the company's principles can be found at https://ai.google/principles/), Microsoft (the company's principles can be found at https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6) or Tesla (the company's principles can be found at https://www.tesla.com/AI), just to name a few [last accessed April 18, 2022].

193 Some of these, and other, statements of ethical principles issued by institutions and companies, can be found in Greene, Hoffmann, & Stark, 2019, Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. The authors analyze diverse statements and conclude they are powerful instruments to build a shared ethical frame and an ethical dialogue. However,

> «[…] this frame is not an innocuous one; rather, it sets and shapes the 'moral background' that make conversations around ethics and technology possible in the first place. Specifically, it offers a deterministic vision of AI/ML, the ethics of which are best addressed through certain kinds of technical and design expertise.» (Greene, Hoffmann, & Stark, op.cit., p. 2129).

[194] Values in Design (ViD). The relationship between human values and technology design that include «[…] Value-Sensitive Design (VSD), Values @ Play, reflective design, adversarial design, and critical technical practice.» (Greene, Hoffmann, & Stark, op. cit., p. 2123).

[195] The Toronto Declaration. Protecting the right to equality and non-discrimination in machine learning systems. The text of the declaration is available at https://www.torontodeclaration.org/ [last accessed April 19, 2022].

[196] The Montreal Declaration for Responsible Development of AI can be found at https://recherche.umontreal.ca/english/strategic-initiatives/montreal-declaration-for-a-responsible-ai/ [last accessed April 19, 2022].

[197] The Asimolar AI Principles can be found at https://futureoflife.org/2017/08/11/ai-principles/ [last accessed April 19, 2022].

[198] The Barcelona declaration for the proper development and usage of artificial intelligence in Europe can be found at https://www.iiia.csic.es/barcelonadeclaration/ [last accessed April 19, 2022].

[199] The Guidelines can be found at http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf [last accessed April 19, 2022].

[200] The UNI Global Union's 10 Principles for Ethical AI can be found at http://www.thefutureworldofwork.org/opinions/10-principles-for-ethical-ai/ [last accessed April 19, 2022].

[201] European Group on Ethics in Science and New Technologies. (2018, March 9). Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems. Retrieved March 22, 2022, from Research and Innovation: https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1.

[202] UNESCO. (2021, November 23). Recommendation on the Ethics of Artificial Intelligence. Retrieved April 1, 2022, from https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000381137&-file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_e86c4b5d-5af9-4e15-be60-82f1a09956fd%3F_%3D381137eng.pdf&locale=en&multi=true&ark=/ark:/48223/p.

[203] Besides that, trust is a dynamic process and includes initial trust and formation and continuous trust. Siau, K., & Wang, W. (2018). Building Trust in Artificial Intelligence,. Cutter Business Technology Journal, 31(2), 47-53. Retrieved April 16, 2022, from https://www.researchgate.net/profile/Keng-Siau-2/publication/324006061_Building_Trust_in_Artificial_Intelligence_Machine_Learning_and_Robotics/links/5ab8744baca2722b-97cf9d33/Building-Trust-in-Artificial-Intelligence-Machine-Learning-and-Robotics.pdf, p. 47.

[204] Ibid, p. 50.

[205] For instance, Robinson, S. C. (2020, November). Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI) . Technology in society, 63. doi: https://doi.org/10.1016/j.techsoc.2020.101421, explored policy AI documents from Norway, Denmark Finland and Sweden and analysed how cultural values – such as trust, transparency, and openness were, or not, upheld in the documents.

[206] Huang, H.-Y., & Bashir, M. (n.d.). Users' Trust in Automation: A Cultural Perspective. (J. Chen, Ed.) Advances in Intelligent Systems and Computing , 595, 282-289. doi: https://doi.org/10.1007/978-3-319-60384-1_27.

[207] Kim, et al., 2021, When is it permissible for artificial intelligence to lie?: A trust-based approach.

[208] The authors confronted Scandinavian countries, and the US.

[209] Kim, T. W., Lu, T., Lee, K., Cheng, Z., Tang, Y., & Hooker, J. (2021, March 16). When is it permissible for artificial intelligence to lie?: A trust-based approach. Ethics of Conversational AI. Retrieved April 18, 2022, from https://arxiv.org/pdf/2103.05434.pdf, p. 6.

[210] European Union Agency for Fundamental Rights. (2020). Getting the Future Right. Artificial Intelligence and Fundamental Rights. Retrieved April 16, 2022, from https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf, p. 49.

[211] See Singler, B. (2017). An Introduction to Artificial Intelligence and Religion For the Religious Studies Scholar. The Faraday Institute for Science and Religion, 215-231. doi:10.1558/imre.35901.

[212] Alternatively, create new religions such as «Data-ism» that Yuval Harari presented (Harari, Y. (2017). Homo Deus: A Brief History of Tomorrow. Random House).

[213] In another perspective, AI can also contribute to Religious Studies. See, for example, Reed, op. cit.

[214] For instance, and just to state an example about the role of the Catholic Church in the challenges presented to humankind, the Catholic Church reaction to the industrial revolution and the Communist Manifesto from

Marx and Engels. The Church reacted through Pope Leo XII and His Encyclical Rerum Novarum. (Almeida Araújo, M., & Meireis, A. (2019). Understanding Dignity (of Workers) across the Catholic Thought. Minerva: Revista de Estudos Laborais, 27-62. Retrieved from https://www.researchgate.net/publication/339845373_UNDERSTANDING_DIGNITY_OF_WORKERS_ACROSS_THE_CATHOLIC_THOUGHT, p. 37). For a developed study about the topic see Moreira, A. J. (2021). Escravidão, Dignidade e Trabalho. Coimbra: Almedina.

Today AI also presents new challenges to labour and employment and is an imperative to the «*post-work future*», using Singler's expression (Singler, op. cit., p. 221).

[215] Robinson, op. cit., p. 13.

[216] The High-level expert group on AI was appointed by the European Commission and, besides the Ethics Guidelines for Trustworthy AI (Independent High-Level Expert Group on Artificial Intelligence to the European Commission, 2019), the group also delivered the Policy and Investment Recommendations for Trustworthy AI (Independent High-Level Expert Group on Artificial Intelligence to the European Commission, 2019), the final Assessment List for Trustworthy AI (ALTAI) (Independent Hugh-Level Expert Group on Artificial Intelligence, 2020) and a Sectoral Considerations on the Policy and Investment Recommendations (Independent High-Level Expert Group on Artificial Intelligence, 2020).

[217] Independent High-Level Expert Group on Artificial Intelligence to the European Commission, 2019, p. 5.

[218] To sum it, is has to be «[…] fair and impartial; robust and reliable; respectful of privacy; safe and secure; responsible and accountable; and transparent and explainable.» (Ammanath, op. cit., p. 9).

[219] Independent High-Level Expert Group on Artificial Intelligence to the European Commission, 2019.

[220] See AI: examples of ethical dilemmas. Available at https://en.unesco.org/artificial-intelligence/ethics/cases [last accessed April 19, 2022].

[221] The performance of AI machines depends on the data, but the way data is collected, and classified is also culture-specific. World Commission in the Ethics of Scientific Knowledge and Technology, 2019, Preliminary Study on the Ethics of Artificial Intelligence, p. 7.

[222] Algorithms are modelled and produced in the social world and, so, they should not be seen outside their context. «Their existence and design are a product of social forces, as are their implementations and redesigns» Beer, D. (2017). The social power of algorithms. Information, Communication & Society , 20:1, 1-13. doi: https://doi.org/10.1080/1369118X.2016.1216147, p. 4.

[223] UNESCO, 2021, Recommendation on the Ethics of Artificial Intelligence, p. 23.

[224] Ammanath, op. cit., p. 164.

[225] Enhance digital skills and competences for the digital transformation requires
> «[…] basic digital skills and competences from an early age digital literacy, including tackling disinformation computing education good knowledge and understanding of data-intensive technologies, such as artificial intelligence (AI) advanced digital skills, which produce more digital specialists ensuring that girls and young women are equally represented in digital studies and careers». (European Commission. (2021). *Digital Education Action Plan (2021-2027). Retrieved April 7, 2022, from https://education.ec.europa.eu/focus-topics/digital-education/about/digital-education-action-plan*).

[226] European Commission, 2016, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions a New Skills Agenda for Europe (COM(2016) 381 final).

[227] The impact of AI in labour is quite significant and, if not addressed properly it might exacerbate inequalities considering the different level of digital skills, especially in lower skills jobs (Nedelkoska, L., & Quintini, G. (2018). Automation, skills use and training. OECD Social, Employment and Migration Working Papers, 202. doi:https://doi.org/10.1787/2e2f4eea-en)

[228] European Commision, 2021, Digital Education Action Plan (2021-2027).

[229] About the topic see James A. , 2020, Planning for Mass Unemployment: Precautionary Basic Income.

[230] The Commission supports research into human-AI field, European Commission, 2018, Communication from the Commission on Artificial Intelligence for Europe COM(2018) 237 final, Point 3.2 of the communication «Preparing for socioeconomic changes».

[231] Ebers, M. (2021, July 15). Regulating Explainable AI in the European Union. An Overview of the Current Legal Framework(s). In L. Colonna, & S. Greenstein, Nordic Yearbook of Law and Informatics 2020: Law in the Era of Artificial Intelligence. Retrieved May 1, 2022, from in: Liane Colonna/Stanley Greenstein (eds.), Nordic Yearbook of Law and Informatics: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3901732).

[232] The different «[…] scale and the power generated by AI technology accentuates the asymmetry between individuals, groups and nations.» World Commission in the Ethics of Scientific Knowledge and Technology, 2019, p. 8.

[233] «A computational tool in which you understand the input and output but not the underlying procedure.» (Barrat, op. cit., p. 75).

[234] OECD Council, 2019, Recommendation on Artificial Intelligence.

[235] Independent Hugh-Level Expert Group on Artificial Intelligence, 2020, pp. 14-15.

[236] Committee of Ministers, 2019, Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes.

[237] Leyen, 2019, Political Guidelines for the next European Commission 2019-2024, p. 14.

[238] The AI index Report from 2019 showed the trends in AI education (Artificial Intelligence Index Report 2019. (2019). Stanford University. Retrieved May 1, 2022, from https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf)

[239] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016

[240] About the topic see Nemitz, op. cit..

[241] Moreover, other legislative initiatives, like the US Act.

[242] Vera Lúcia Raposo is her assessment of the European Draft act on AI addresses the issue of the balance between protection of HR and the price for innovation (Raposo, V. L. (2022, March 18). Ex machina: preliminary critical assessment of the European Draft Act on artificial intelligence. International Journal of Law and Information Technology, 30 (1), pp. 88-109. doi: https://doi.org/10.1093/ijlit/eaac007)