Artificial Intelligence and Human Rights, an Unequal Struggle

Maria Stefania Cataleta, Anna Cataleta

CIFILE Journal of International Law (2020), *Journal Vol. 1, No. 2, 40-63 (2020)*

Online ISSN: 2563-3341

Artificial Intelligence and Human Rights, an Unequal Struggle

Maria Stefania Cataleta *1* ***Anna Cataleta*** *2*

1. *Ph.D., LLM, lawyer and researcher at the University of Côte d'Azur, LADIE, Email: Franceavvmscataleta@tiscali.it*
2. *Lawyer and lecturer Politecnico, Milan, Italy, Email: annacata@libero.it*

DOI: 10.30489/CIFJ.2020.223561.1015

ARTICLE INFO

Article history:
Received: 14 March 2020
Accepted: 12 May 2020
Online: 22 June 2020

Keywords:
Artificial intelligence robotics, human rights, privacy discrimination

ABSTRACT

Artificial Intelligence (AI) is a kind of intelligence that was born in the 1950s and is an integral part of the digital revolution. Progress made by AI has permitted the birth of systems capable of rivalling human capacities or, in some cases, surpassing them. The progress of the intellectual capacities of AI will change the way of life for human beings and will revolutionise the world of employment. Intelligent systems present problems regarding individual rights and responsibilities, because as technology replaces more and more of what humans have typically done, our individual roles will become more blurred. The goal of this analysis is to measure the developments of AI in relation to its impact on society, in particular on human rights, fundamental liberties, and ethics. This is an unexplored topic within the vast field of AI upon which this paper will expound.

INTRODUCTION

In 1997, the programme Deep Blue, developed by researchers of Carnegie Mellon University, defeated the world champion of chess Garry Kasparov. The winner did not accept the defeat, under the assumption that a machine could never compute such intelligent strategies and accused scientists of rigging the match. This victory started the discussion on the primacy of the human mind over devices. Chess had long been considered a strategic endeavour representing the peak of human cunning, strategy and intelligence; this upset was groundbreaking. Nowadays, programs capable of playing chess are so powerful that they are no longer used against human adversaries, but against computers[1].

This is what we call Artificial Intelligence, a kind of intelligence that was born in the 1950s and has proven to be an integral part of the digital revolution. This revolution has been part of a myriad of social transformations[2].

Nowadays, the progress made by A.I., alongside robotics, has given way to systems capable of rivalling human capacities, or in some cases, surpassing them. These systems are mostly autonomous because they are capable of learning from their own experiences through machine learning, and will continually improve themselves through the digestion of ever-larger stores of data. Furthermore, they are capable of doing things that they were not programmed to do; computers no longer do *only* what they are programmed to do.

The progress of the intellectual capacities of A.I. is changing the way human beings and will revolutionize the world of employment, along with many sectors of human life. There are systems capable of free thoughts and actions which complicate legal ideas around who is the subject legally responsible for a given activity: the computer? The algorithm? The computer scientist? The user? Furthermore, intelligent systems create problems such as understanding if they have individual rights and responsibilities and knowing which actions they can pursue without violating fundamental human rights.

We now find ourselves in a digital economy that has A.I. at its forefront. The largest companies of today, the ones that run our daily productivity, are all based on A.I.:

Facebook, Google, Amazon, Microsoft, Baidu, Alibaba and Tecent.

These companies flirt with A.I. transhumanist projects that will give unlimited powers to man, such as modifying his genome, reprogramming his brain and neutralizing death. Doas AI make men similar to gods?

The goal of this analysis is to calibrate the developments of A.I. concerning its impact on society, in particular on human rights and fundamental liberties, because algorithms, especially machine learning algorithms, are often accused of propagating inequity, discrimination and opacity. This is an area in which, only recently, ethical legislative regulation has begun to lay its foundation.

Are some rights destined to die among the spread of A.I. or will human intelligence be able to identify some boundaries capable of avoiding the dangers that inevitably accompany the digital evolution?

1. The rivalry between A.I. and human intelligence

Speaking of "creative disruption," the economist Joseph Schumpeter indicated the technological steps involving the disappearance of specific activities and the birth of other more creative solutions. This principle has held from the first industrial revolution to the information-revolution. Three great revolutions have characterized the last two centuries: the first, from 1770 to 1850, with the first factories, the steam engine and the railway network; the second, from 1870 to 1910, with the birth of aviation, automobiles, electricity and telephones; the third, beginning around 2000, with the arrival of the N.B.I.C. (Nanotechnologies, Biotechnologies, Informatics and Cognitive Sciences), for which life operates on the nanometric scale, one billionth of a meter[3]. N.B.I.C. develops exponentially and throw our society into rapid evolution with unforeseen consequences. N.B.I.C. is quickly becoming grouped in the form of a new type of science of the XXI century, one structured on the Internet of things and A.I.

Many definitions of A.I. have been offered, the first of which came in 1956 during the Dartmouth Summer Research Project on A.I. John McCarthy, one of the founding fathers of the discipline, defined "intelligent" any system capable of performing actions that would be qualified as intelligent if a human being accomplished them. On this occasion, it was presented the first program explicitly projected to emulate the human capacities of problem-solving. The conference demonstrated how machines could use language, formulate abstractions and concepts, resolve problems about humans and better themselves. McCarthy was convinced that computers could simulate many of the cognitive functions of human beings. Thus the expression A.I. was invented, distinguishing it from simple automation, where each process that could be carried out, automated by a specific algorithm [4].

The result of the conference led to the common definition of A.I.:

a set of scientific methods, theories and techniques whose aim is to reproduce, by a machine, the cognitive abilities of human beings. Current developments seek to have machines perform complex tasks previously carried out by humans.

Essentially, the result of an operation performed by an intelligent system is not distinguishable from a process carried out by a human[5]. In other words, it is a discipline that studies the design, the development and the realization of systems capable of simulating human ability, reasoning and behaviour.

Moreover, it is arduous to equal human intelligence from artificial intelligence, neither the speed of calculation can be only an indicator of the fact that machines possess superior intelligence. Human capacities, in other words, are not a good metre of judgement for A.I., mainly if you can consider that there are activities of A.I. that are precluded from human intelligence and that notwithstanding express intelligence. A tsunami alert system is not comparable with human abilities because it is based on minimum movements of ocean heights undetectable by human senses, those incapable of perceiving the submarine geological upheavals. However, the study of the human mind is still concentrated on understanding how it is possible that neurons, masses of homogeneous cells, through their interconnections (the

synapsis) can change electrical or chemical signals and under this, perform the most varied activities.

Something similar happens in the research of artificial neural networks, where one can search for understanding how to make interconnections converge toward the most acceptable solution in a reasonable time. Artificial neural networks, more than learning, seem more like imitators of strategies taken from a large number of examples. A.I. learns by doing by way of machine learning, of which one of the principal aspects is deep learning. Deep learning refers to the use of artificial neural networks with many inner layers, called "hidden layers." Deep learning is a system of education and classification that, across networks of artificial digital neurons, allows a computer to acquire some capacities of the human brain. The artificial neural network seems to be closer to the social nervous system. In the MLP-Multi-Layers Perceptron networks, there are hidden neural layers where every neuron in a level is linked to all neurons of the immediately preceding level and of the level right after. The real power of the algorithm is given by the capacity to train the neural network and to allow it to gain experience.

It is with deep learning that A.I. was born, whose name was coined about "smart" calculations, similar to those of a powerful calculator, capabilities much less complicated and intelligent than their current uses in recognizing the contents of images or understanding spoken language[6].

Many problems that at first sight would seem to require logic and reasoning can be resolved through machine learning. Self-learning of computers has an exponential tendency; they continuously improve and refine output by accumulating experiences. Machine learning is defined as the mother of all the algorithms in A.I., with all its variables of self-learning. Firstly there is supervised learning: data associated with information that interests us is given to the algorithm, on the basis that the algorithm will learn how to understand and how to behave (an example is the classification of potential clients on the base of the profile and of the history of buying of other customers). Then, there is unsupervised-learning, where the algorithm is more complex because it needs to extract still unknown information from the data, or further reinforce its own learning, in which the algorithm has a goal to reach and thus auto-defines a way in which it behaves which can change in the face of differing situations.

Scholar Jerry Kaplan believes that a computer can be more intelligent than a human being, even if limited in manner[7]. It is true that computers, in a wide range of intellectual duties, are superior to man, but this does not necessarily mean that it will dominate us.

Furthermore, there is a question which scholars try to answer when investigating the supremacy of machines over man: if they are equipped with a mind and thoughts. In this field, two theories face each other, that of "strong" A.I., referring to machines which have a mind or in any case will end up having one someday, and that of "weak" A.I., which considers that machine realities are only simple simulations and not a duplication of real intelligence. The conceptual crossroads are between the possibility that machines can be truly intelligent or merely capable of acting like they are[8].

According to the distinction made by American philosopher John Searle, the weak A.I. acts and thinks as if it had a brain, but it is not intelligent; it limits itself to emulate the human brain. To offer the best answer to a problem, it investigates similar cases, it studies them and chooses the most rational solution. The weak A.I. does not understand all the human cognitive processes but only deals with problem-solving. That is, it answers problems based on known rules. The strong A.I., instead, has cognitive capacities indistinguishable from human capabilities, but according to Searle, we are still aways away from this reality.

"Expert systems" are highlighted in this context: software that reproduces performance and the knowledge of experts in specific fields. The inferential engine is at the centre of these systems. In other words, it deals with an algorithm which, similarly to how the human mind works, starts from a proposition whose truth derives from the content of the first proposition, according to deductive or inductive logic.

Conforming to the Turing test, though would be extended to the machines and this would have happened around the end of the XX century, whereas skeptics, like John Searle, sustain that machines cannot think at all, because it is an exquisitely human activity, that computers are limited in

simulating. Nevertheless, the actual capacities of machines put their beliefs on the superiority of the human mind to the test.

Kaplan analyzes the issue from the perspective of free will and of the capacity of machines (following anthropomorphic criteria) to make decisions, the same as a human would. Devices in their decision making would be capable of applying knowledge and be competent in assuming risks and modifying plans based on added information: using analogies to solve concrete cases. No motive has been found to believe that man and machine follow different principles regarding decisional processing[9].

2. The decisional capacities of the machines and their impact on society

The decisional capacity of machines presents another problem: making autonomous decisions requires responsibility. Whenever one signs a contract or makes a purchase online, that agreement is binding even if there is not a human being as a counterpart. Behind these contracts, there is always a company. Nevertheless, this will probably change because an ever-increasing number of "intelligent agents" are always posed in the autonomous modality concerning those who they should represent. In the case of online purchases, the responsibilities would lie outside of the human being for the action taken autonomously by an intelligent agent. There are circumstances where the presence of a human being is problematic because he may not want to accept the responsibility of the actions of autonomous smart agents. After all, the consequences of those actions escape the human control for which he has no role.

The Resolution of the European Parliament of 16 February 2017 lays down recommendations for the European Commission concerning norms of civil rights regarding robotics[10]. The Parliament indicated as a general theme that in A.I. and robotic technology, human capacity should be integrated into technology rather than substituted. Concerning the responsibilities of intelligent agents, there are problems regarding the imputability of duties, which generally fall on the agent subject; but can the intelligent system set up its autonomy in a natural way? A.I. raises essential questions surrounding the principles of responsibility, whether they fall upon the manufacturers, the programmers who created the decisional algorithms, or the users of the program, in terms of *culpa in vigilando*, which can be invoked the moment in which a user was to recognize any wrong decision by an intelligent system[11].

Company law could be used as a model to assign rights and responsibilities in the field of A.I. Under company law, companies specifically limit the individual accountability of stakeholders and shareholders to exclude personal liability, which otherwise could have legal or fiscal ramifications, to maintain professional activities. Similarly, the creators and producers of A.I. could remain immune from the legal responsibilities of the actions taken by their created device. Yet, differently from companies that depend entirely on human beings to endorse every step, an intelligent system can simply act singularly, which is precisely what makes it particularly dangerous.

The conviction of the primacy of man over the machine would lead one to think that there is always someone behind a mechanism that controls it. But as we have seen, it is not still so. An artificial intelligence system can commit crimes, damaging the social order, hurting public interest, as well as harming individuals all on its own[12]. Regarding human and machine interaction, there is an important distinction to be made under current law regarding the delegation of processes and the delegation of decisions. Currently, *humans* delegate the choices of machines: if one wants to wash his clothes, it is a *man* deciding to use a washing machine even though he charges the *process* to the computer. It remains the *decision* of man to determine when and how to compute an action, also if a machine performs such work. Devices with decisional capacities pose risks because they act concerning a system of values, expressed by algorithms, that are not only non-universal but can even be in conflict between themselves[13].

One of the most noted applications of A.I. is that of self-driving cars, which is estimated to reach an adoption rate of 75% by 2035. This use of A.I. has raised ethical issues which have been subject to much research. For example, in 2018, M.I.T. researchers published in their journal, *Nature*, the results of the survey *The Moral Machine experiment*, that involved 2 million people in 233 countries. The study was aimed at understanding people's opinions on the

choices that a self-driving car should make in an emergency.

An autonomously driving car could be presented with having to make moral decisions in extreme situations. For example, if two types of accidents could occur, which would implicate either the death of an older adult or the end of a child, the car would have to decide for itself which path to take. In such cases, it is crucial to program the machines to be able to understand how to make moral decisions, following an area of research called computational ethics, which aims at creating moral agents. Such issues will become more and more present as A.I. systems and human beings form closer relationships ever. The technological challenge will be, therefore, to create programs capable of behaving "adequately" in society and concerning legal and/or ethical norms that will act as a rubric for the behaviour of these intelligent systems.

The spread of A.I. systems raises concerns on how to mitigate the impact of machine intrusiveness, enabled with decisional capabilities within our society. A.I. will create autonomous technologies with behavioural rules based on data analysis, which is based on statistics, not morality. Furthermore, A.I. is increasingly able, often more than a human being, to do complex duties and to solve articulated problems by choosing between several possible alternatives. A.I. seems bound to substitute not only manual labour but also intellectual labour[14].

There are, however, many concerns from academia[15]. The first is that an individual can lose control of his own life in that autonomous technologies could reduce human involvement in society to a controversial stance. The primacy of man falters in the face of the decisional capacities of machines because rational decisions could soon become dependent upon the outputs of automated intelligent systems.

Among the concerns, there is the motive behind A.I. application, whether it is for net-good optimization or rather profit and power. Furthermore, because the personality of A.I. is directly related to that of who programs it, we could end up with malignant machines based on a psychopathic A.I.

Another problem is the risk of social discomfort that could come about from the increase in automation of manual labour jobs, *blue-collar jobs*, or intellectual labour jobs, *white-collar jobs*, or only due to increased cognitive dependence upon A.I., in an age where most knowledge comes from the web. Not to mention the last scare linked to the use of A.I. according to criminal modalities (*cybercrime*), or by the point of view of war (*cyberwarfare*)[16]. That is what some scholars intend with the expression "malicious A.I.": the criminal employment of A.I. to undermine society's security, so much to make you think of the XXI century as an age in which A.I. will be at odds with humans[17].

In preparation, scholars have already started to formulate ways in which to combat the potential overextension of A.I. First of all, it is essential to increase and to improve international collaboration between different stakeholders to reach a shared vision on how to promote an advantageous development between society and A.I. Secondly, it is necessary to encourage growth in politics, which directions the evolution of A.I. toward improving living conditions and the common good, following ethical norms. Lastly, it is necessary to avoid the obsolescence of anthropogenic knowledge: orientating the economic, political and educational system to improve individual capacities so that they are always at the forefront in terms of autonomous intelligent technologies[18].

According to research by McKinsey Global Institute (M.G.I.) (*Modeling the Impact of A.I. on the World Economy*), A.I. could increase the global economy by around 13 trillion dollars by 2030, with a growth of about 1,2% of the Gross Domestic Product (G.D.P.) each year. If this prediction were to be confirmed, the impact would be similar to that produced by the steam engine in the 1800s or robots in industrial production in the 1990s. Indeed, A.I. is predicted to be a disruptive technology. Yet how exactly its force will impact society is still to be understood. Scholars sustain that the impact of A.I. will not be linear and will increase according to the classic model of the S curve[19]. Yet, despite rapid adoption, economic acceleration will depend on policies espoused by different countries based on support, acceptance, or even hostility toward these technologies through legislation[20]. The impact of A.I. on society remains a mystery.

3. The vulnerability of human rights under the pressures of A.I., the international scenario

Artificial Intelligence has proven to be fundamental for our society: it is already commonplace, even if we do not explicitly recognize it. These systems will continue to persevere as they take over evermore functions that humans traditionally have done. As we delegate more responsibility to machines in regards to autonomous decision making, we must guarantee proper accountability regarding human rights protection.

One such example of the use of A.I. for good yet also use, which violates human rights, is one of social monitoring. Currently, governments use video surveillance and biometric techniques combined with A.I. to track and monitor terrorists. These computer systems are trained to identify persons of interest and follow them. Yet, this technology, when used on ordinary citizens, is a clear violation of the fundamental right to privacy.

Where international terrorism is concerned, governments require and implement many new technologies, such as video surveillance and biometric tracking, to thwart illegal and threatening behaviour. These government activities make our lives more secure and do work to hinder criminal activities. However, these same technologies actively monitor and track ordinary citizens, which presents a violation of individual privacy and could entail future discriminations based on religious beliefs, health conditions, or even political opinions. Also, evolution within the nanotechnology sector raises even further problems. To develop these kinds of technologies, there will be the inherent, unknown involvement of random, third party citizens. Their rights must be considered within the principles of solidarity and social justice[21].

In the face of technological and scientific progress, the concept of a *legal person* is pushed to its limits, for *"the scientific and technological world, artificial in conceptual nature, come to encroach upon the already defined legal dimension of a person, an artificial concept in itself"*[22].

On the ongoing path of social legitimization of technological progress, human rights represent a referential normative principle. This legitimization *"cannot be accepted only on the grounds of security or on the logic of economic efficiency"* and *"must always remain measured by the metre of democracy and respect of people"*[23].

Along with this technological development, the underlying concept of all human rights is called into question: that of dignity. It is a notion that is at the base of all human rights and of the Natural Equality of human beings, as such, it is present in the United Nations Charter and of the Universal Declaration of Human Rights of 1948, in particular in the preamble (*"Whereas recognition of the inherent dignity and of equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world"*) and in articles 1 (*"All human beings are born free and equal in dignity and rights [...]"*) and 2 (*"Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind [...]"*). Dignity makes human beings entitled to inalienable rights, ones which guarantee natural equality, protecting us from any form of discrimination[24]. Technology risks to undermine this equality[25].

The problem of equality is linked to that of no discrimination, two specular concepts that represent positive and negative articulations of one unique principle. Indeed, equality means to treat all cases equally, and no discrimination serves to prohibit the biased treatment for all reasonably motivated cases. All treaties on human rights provide the principle of equality[26]. For certain serious discriminations, such as those concerning race, ethnic origin, sex or religion, stringent tests are posed upon the State to justify their existence.

The conventions on human rights prohibit direct discrimination, which occurs when a person is treated in a disadvantaged way concerning another who is in a similar situation, and that of indirect discrimination, when a person who is formally treated like the others, suffers a disadvantage from a predefined equal treatment. Now, treaties on human rights do not ask for discriminatory intent, as they currently also prohibit unintentional discrimination[27].

It is true, as the U.N. Committee on human rights observed, that not all differentiations in treatment constitute discrimination. When founded on reasonable

and objective criteria, there may have a legitimate goal[28]. Furthermore, the European Convention on Human Rights, in the case concerning the use of language in Belgian institutions, ruled that the principle of equality is compromised if the differentiated treatment does not have any reasonable or objective justification and that the measure that has made the differentiation must pursue a legitimate goal and present a rapport of reasonable proportionality between the means employed and the aim continued. In other words, it is necessary to verify if the measure that makes the differentiation were to pursue a legitimate goal and if it pursues that objective with means proportionate to the purpose.

Where legitimate purpose is missing, justification will also be missing, resulting in illegal discrimination; at the same time, even in the presence of a legitimate purpose, it will be unlawful discrimination if the target is pursued with disproportionate means. It follows that the only case of legal differentiation will happen when the legitimate purpose is pursued with proportionate means. It is a double test that has been accepted in the context of human rights and which is adopted by supervisory bodies such as the Committee of human rights of the United Nations. As a rule, then the burden to prove discrimination is on the victim, whereas the burden of proving the presence of a cause of justification lies in the State[29].

Having discussed the principle of equality and non-discrimination, one can ascertain that there also exist "algorithmic prejudices" or *bias*, able to cause social discrimination, which the increase of available data and individual computing capacities of A.I. systems risk amplifying. Until the risks of this type are delineated, it is crucial to develop an *ethics of data*[30]. Aimed at this, the European Union is preparing to publish the first draft of an Ethical Code, on the base that A.I. must never damage the dignity, the physical security, the psychological security, nor the financial security of human beings, animals or nature.

In December 2018, a group of experts drew up the "Draft Ethics Guidelines for Trustworthy A.I[31]. With this document, the European Commission warned of the risks associated with A.I., despite its considerable advantages, and recognizes the need for an anthropocentric approach to A.I. This is the only approach capable of guaranteeing the dignity and autonomy of people, who must always be given the power to supervise machines[32]. Even the Council of Europe recently warned against the risk of "social discrimination" provoked by algorithms.

Take as an example, two sectors to describe such risks: the system of facial recognition, and that of justice. Numerous studies ("Gender shades" by the researcher Joy Buolamwini from M.I.T. included) sustain that facial recognition can jeopardize our freedoms. Indeed, research conducted on different facial recognition systems (such as I.B.M. Watson, Microsoft Cognitive Services and Face++) has shown that some ethnicities are treated in a more imprecise way in respect to others. Notably, precision identification for white men was 99%, but only 34% for women with a dark complexion. This is because algorithms of these systems are based on subject-data-inputs, which are prevalently male and of light complexion. It is evident that mistakes in programming algorithms have been committed, and it is not easy to correct them either. This is due to the quantity of data analyzed by the algorithms, which grows exponentially, creating errors that are deeply buried inside the artificial neuronal layers.

Another study, this time produced in 2018 by the A.C.L.U. (American Civil Liberties Union) an American Association in defence of civil rights, using Rekognition, analyzed the photos of American parliamentarians in a database of about 25 thousand images and demonstrated that in 5 percent of cases, an inexistent correspondence emerged between parliamentarians and criminals.

However, what makes matters worse is in the fact that these false positives, 39% of them, concerned deputies of dark skin. Similarly, it occurred that recruitment software for potentially new Amazon employees favoured hiring males instead of females. Paradoxically, technological progress worsens the bias in A.I. systems, because the enlargement of the databases does nothing but automate and standardize the error.

Let us turn to the risks for the legal system[33]. Scholars have offered examples coming from the American legal system where A.I. is used for crime prevention. The programs are developed to calculate the probability that

the accused would be a repeat offender, aiding judges in establishing appropriate sentences to avoid such risk.

Well, it has been shown that the *ab origine* collection of discriminatory data, such as the mapping of certain urban areas or the collection of data of potential criminals or victims, is able to consolidate prejudices, to the detriment of rights and fundamental liberties. It has been opportunely observed that to entrust a judgement on a crime to an algorithm based on the possibility that a future crime could occur is an obstruction of proper legal discipline[34].

It is precisely in this area, specifically in the field of predictive justice, that there is a risk of massive violation of human rights through the use of AI-based devices, such as the spread of risk assessments tools (used in the United States) or computational tools based on A.I. capable of calculating the probability that a person will evade trial or commit crimes. These are mechanisms that examine a large number of data related to the past, such as socio-economic or family status and other factors and identify patterns, i.e. recurrences, based on a more reliable statistical basis than that based on human judgment. Risk assessments are used mainly in North America at all stages of judiciary processes, from the preliminary stage, where the release of the suspect must be assessed, to the decision-making stage[35.]

A well-known tool is the "*Correctional Offender Management Profiling for Alternative Sanctions*" (C.O.M.P.A.S.), an algorithm that analyzes the answers to a questionnaire of 137 questions related to criminal involvement, relationships/lifestyles, personality/attitudes, family, and social exclusion[36], one which has been the subject of harsh criticism because it produces discrimination based on race, creating unequal treatment disadvantageous to individuals of colour. Similarly, it creates bias related to the probability of committing crimes, which disproportionately affects individuals of colour twice as much as individuals of lighter complexion. To eliminate the discriminatory effects of C.O.M.P.A.S., the Laura and John Arnold Foundation has created another tool, the Public Safety Assessment (P.S.A.), which would eliminate the negative impact of information concerning gender, race or economic conditions. It is a tool that can assess, on the basis of nine risk factors, whether an individual will appear at trial and commit an offence if he or she is released before the trial[37]. It would reduce the risk of bias because the number of criminal convictions would have a greater influence than other assessments and criteria, because it would be neutral in relation to race and because it would give the last word to the judge and not to the algorithm. In Kentucky, however, through the Administrative Pretrial Release Program, the use of the P.S.A. allows the defendant to obtain the release without the intervention of the judge and without bail, the bail for release, but only based on the use of the tool.

The use of risk assessment tools has also involved the English judicial system, which uses the Harm Assessment Risk Tool (HARM) system for predictive assessments aimed at reducing the risk of recidivism, not free from criticism in terms of violation of privacy, as it takes into account 34 variables, including those related to criminal records, age, gender and postcodes of residence of the individual[38].

In the current state of technological evolution, it is arduous to eliminate the distortions of A.I. systems; for this, it is important to be conscious and to adopt the appropriate precautions. In this regard, Google has implemented its company policy, prohibiting the development of A.I. systems that can transform into tools of surveillance of users.

Therefore, there have been numerous requests from scholars for the rapid development of an "ethic of data," precisely what Europe is undertaking through the European Commission body. Similarly, the United States, through the governmental agency of the Development of Defence of the United States (DARPA), which has the duty of developing new technologies for military use, has been developing tools to instill ethical norms in A.I. machines, through a two billion dollar program.

In order to avoid other scandals such as Cambridge Analytica[39], the future of A.I. largely depends on the ability to solve the problems that are inherent to the increase of data available to the machines and their calculation capacity. The big names of Silicon Valley are already working to reduce these risks linked to the prejudices that are hidden inside A.I. systems[40].

Yet, the problem remains that, when fundamental rights came into play, it is difficult to entrust them to the decision

of an algorithm and skip the computer science model, because discretionary and ethical assessments typical of the human element are paramount.

4. A.I. and the protection of human rights on the European scene, and its countermeasures

During the works of the Conference "Governing the Game Changer-Impacts of artificial intelligence development on human rights, democracy and the rule of law," held in Helsinki in 2019 and organized by the Council of Europe and by the Finnish Presidency of the Committee of Ministers, there was a lively discussion on the impact of the development of A.I. on human rights and the necessity of major research along with trust and transparency in this field.

The conference concluded with the formation of guidelines on how to proceed in the development of A.I. while guaranteeing security and benefits for all. Between the prefixed proposal, timely and thoughtful policy responses were called for, which could be put among the priorities of governmental political agendas. An invitation was addressed to all States and parties involved to coordinate such initiatives and share information and good practices. It has been underlined that A.I. should be developed in a way to put human beings at the center of advantages for people and society. The need to establish efficient mechanisms of supervision and structures of democratic vigilance in relation to the design, the development and the implementation of A.I. and the necessity to acquire the public consciousness of potential risks and advantages of A.I. were the main points highlighted.

The urgency of efficient and legitimate mechanisms to prevent the violation of human rights, discriminations, inequalities and prejudices have become apparent. It has been recognized that the transparencies of algorithms is crucial for the creation of trust and to guarantee the due protection of rights while at the same time agreeing that equality in front of law should not be compromised by the calculation of algorithms. It is hoped that A.I. is used in a way to guarantee that technological progress will happen in accordance with the principles of human rights, democracy and the rule of law and with respect to existing international instruments of reference. It has been urged that the Council of Europe will continue to develop recommendations, guidelines and specific codes of conduct for the sector in order to promote human rights and democratic processes, even though monitoring the impact of A.I. on the common basis of democratic societies.

These are all good proposals, but in order to reach these goals, it is necessary to regulate A.I. and guarantee its development with legal principles, and not leave out ethical-social aspects to which A.I. must comply.

Moreover, there are the questions of the Council of Europe, an international body with a broader horizon than that of the E.U., which puts into question the adoption of the guidelines based, among others, on Convention 108 and on the Ethical Charter on the use of A.I. in judiciary systems, while, at the same time, taking into account the European Convention for the protection of human rights and fundamental liberties.

Regarding the ethics of the data, the European Commission has established seven requirements for ethical A.I. which industry, research institutes and public authorities must respect. These are human agency and oversight, technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental wellbeing, accountability[41].

The study published in 2018 entitled "Algorithms and Human Rights-Study" on the human rights dimension of automated data processing techniques and possible *regulatory implications* was the basis for the European Ethical Charter regarding the use of A.I. in the judicial system, adopted by the Commission for the Efficiency of Justice (C.E.P.E.J.). The concern, in fact, was that the use of A.I. in this field would not violate the right to a judge and the right to a fair trial through the presumption of innocence, equality of arms and respect for the contradictory, but also the right of non-discrimination, given the use of sensitive data in predictive judgments of responsibility, such as racial or ethnic origin, political opinions, religious or political beliefs, socio-economic conditions, or data related to health or sexual orientation. In this sense, the right to a judge, in accordance with Article 5 of the C.E.D.U., takes on the meaning of the right

to the physical presence of a judge, which, therefore, cannot be replaced by an algorithm[42].

In front of predictive techniques of analysis, which are quite invasive, and discriminatory risks connected to algorithmic choices, the problem of the ethical impact of A.I. models rises. In this regard, there is who exploits the role of data ethics, with all the possible superimpositions between ethical and legal aspects, wherein ethics are called to integrate the law. Ethical evaluations are requested when a model of A.I. compliant with the law intersects the ethical-social values of society. An example could be the management of smart cities through algorithms, where automation poses both ethical and legal questions. In both cases, regulation is necessary.

But a solution is not necessarily a new law. Already existing laws are able to address different legal aspects inherent to these issues, such as that of civil responsibility to the protection of personal data. It is also true that it is necessary to consider that existing regulatory models were formed between the 1970s and 1980s and thus were created under vastly different social contexts, therefore highlighting their current unsuitability. Thus new regulation is necessary, especially the hard law, such as laws and conventions. Unfortunately, these types of legal changes are slow to come about, a clear conflict with the very rapid evolution of technology. In the European environment, an example is the General Data Protection Regulation (E.U.) n. 2016/679, better known as G.D.P.R., adopted on 27 April 2016, published on 4 May 2016 and entered into force on 25 May of the same year and operating from 25 May 2018, but *in nuce* in 2011: one must consider that enactment requires many years. This is alarming considering the fact that within a decade, two entire technological generations can come to pass. The timeframes of regulation cannot keep up with those of technology[43]. It is evident that regulatory systems are always outdated with respect to technological progress. There is no lack of consensus in academia, which has criticized the G.D.P.R. as having an unclear remedy to the risks posed by the algorithms whose "black box″ cannot be opened[44].

In this regard, one should compare the E.U. approach and that of the Council of Europe, who present two different regulatory models whose principles are based on either soft law and hard law, respectively. For example, in the personal data sector, it is possible to quote the Convention 108 of the Council of Europe, whereas, if you look at the E.U. countries, in the same sector there exists the G.D.P.R. and a directive (Directive 2016/680), plus a series of detailed national regulations for the integration of G.D.P.R[45]. What emerges is a varied and complex regulatory landscape[46].

Therefore, the question still exists in the academic field of law regarding A.I. regulation: which solution is best? Regulation, a law, or some different type of instrument within hard law. Complicating these issues, we must acknowledge the fact that A.I. implicates the most varied sectors ranging from military to medical applications. This causes one to wonder how a new, uniform regulation can encompass all these different contexts. The idea proposed is that of adopting a "surgical approach," like in e-commerce, where the legislator, instead of introducing new rules, intervenes on the criticisms which have emerged from the transposition of contractual negotiation in the webspace context. The suggestion is to find a limited number of rules of principles, integrated through sectoral instruments of soft-law[47]. There must be, in other words, an interaction between hard-law and soft-law, and that some authorities are in charge of safeguarding the respect for the law.

The final problem arises when analyzing how to guarantee an A.I. development conforming to legal principles and the problem of creating a judicial body called to oversee the correct application of the regulations in accordance with these principles. Given the polymorphic nature of A.I., that – as we said – includes different sectors, the creation of an independent authority could generate many perplexities regarding its precise areas of operation, and this could feud conflicts with other authorities already operating in different sectors and activities in different countries. It is arduous to think of unifying all of this in a unique subject. Similarly, for the co-regulation, academics have proposed co-decisional models in order to foresee an operative synergy among regulators[48].

5. A.I. and digital security, the protection of personal data online

Machines are capable of creating incredible connections with what we have learned from data input and, from these "self-teachings," can create new information which simulates human behaviour; putting difficult connections regarding human mindspace at our fingertips, connections which a normal man could not arrive at through his own elaboration. By processing the immense quantity of data available today, machines increase human intelligence, and for this, A.I. is sometimes defined as "augmented intelligence."

Every second, billions of internet users give big digital operators fantastic amounts of personal data transmitted over social networks, equating to an annual market value of one trillion dollars. This confirms the Metcalfe law, according to which the value of a network grows exponentially in relation to the number of users. To give an example, every time a user creates a Facebook account, it exponentially raises the value of the network. Thus, the Metcalfe law also applies itself to the added value given to A.I. by every user of a social network. From this immense social, economic and emotional heritage, the big digital operators create the world of A.I. But how is this patrimony of data and the rights that are at its roots protected?

We are speaking, *in primis*, about personal data, because the protection of these data is one of the sectors most involved on a daily basis by the arrival of A.I. systems[49]. The operation of these systems, indeed, is based precisely on elaboration, analysis and treatment of large quantities of information, in particular personal data, data that travel on the net [50]. But, many risks lurk in this same space.

Online mass checks, data theft, phishing and malware are all risks to our digital security, a security that is entitled to a number of rights. First and foremost, there is the right to privacy, which also implicates other rights such as that of expression or freedom of peaceful assembly and association. A number of rights are therefore called into question online. When we transmit information about our movements or habits through a mobile phone, our right to privacy is called into question. When we participate in online public debates, expressing our opinion, we exercise our freedom of expression. When we conduct online searches on a subject of our interest, the right to seek and receive information takes over. Finally, when we use an application to agree to participate in a public demonstration, we exercise our right to peaceful assembly. In all these cases, human rights are at stake: those minimum standards capable of preserving human dignity, which is all so interconnected, non-hierarchically ordered, and interdependent between themselves, implies that the violation of one jeopardizes the enjoyment of others. In the online space, proper and effective digital security ensures the protection of these rights.

Many of our daily online activities are subject to observation or, often, real surveillance. Making a purchase, a reservation, or expressing a "like," are all actions that can provide, more or less consciously, information about ourselves. Today, through the Internet, you can communicate information in countless ways or have easy access to a large amount of data. Every time we put information on the net, these little segments of our lives are brought together to paint a picture of who we are, what our tastes, our beliefs, our movements are, and so on. The diffusion of personal information is not an end in itself, even now the big ones of Silicon Valley have not limited themselves in just using our data to predict human behaviours, but have gone so far as to actually try and modify them. The economic imperatives of giants like Amazon or Facebook erode democracy with their systems, reducing individual awareness, decision-making skills and Internet users' ability to react.

However, not communicating compromising or confidential information online does not automatically protect us from possible violations of our rights, which are trampled on the net every day. Violations can also occur through the use of apparently harmless or irrelevant data. The information is derived from digital data or even more from metadata, specific additional information contextualizing a certain datapoint on which only tangential consent has been given. One can get a lot of information from metadata like interests, political orientations, social life and so on.

In the digital world, therefore, the right to privacy, which is protected by international treaties such as the Universal Declaration of Human Rights (art. 12 *"No one shall be*

subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.") and the International Covenant on Civil and Political Rights, *"No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks"*), in addition to regional instruments such as the Charter of Fundamental Rights of the European Union (art. 7, *"Every individual has the right to respect for his or her private and family life, home and communications"*). But the right to our privacy is constantly undermined by the use of the Internet, which constantly is fed more and more information. These can be provided with our consent but also fraudulently extracted and used by criminal networks to extort us for money, by governments to carry out mass checks or through more mundane ways such as by companies to model their advertisements according to our personal profile.

While digital communication has, on the one hand, revolutionized the world of work and interpersonal relationships, it has, on the other hand, made our privacy more fragile, making it more permeable to violations. As we have said, interference in our right to privacy may involve the violation of other rights, such as the right to freedom of expression, precisely because of the interdependence and interconnection of human rights, so the violation of one of them threatens all others. Freedom of expression is guaranteed, among other documents, by the International Covenant on Civil and Political Rights (art. 19.),

"Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice. The exercise [...] carries with it special duties and responsibilities."

No discrimination on the grounds of nationality, gender, genetic characteristics, ethnic or social origin, religion, language, political opinions, property, disability, age, or sexual orientation is allowed in the exercise of that right, or any other status. Freedom of expression through the web can only be *constrained* by law and in such a way that its limitation is necessary and proportionate to a legitimate purpose, such as, for example, the protection of the national interest.

This right may be violated by mass controls – contrary to both the right to privacy and the freedom of expression - which act as Bentham's Panopticon memory - and which causes the user to start to censor themselves for fear of being watched. In this sense, freedom of expression is restricted because it does not allow the user to express himself freely on the web, which is why freedom of expression is closely linked to the right to privacy in the digital world because if you have the perception that your privacy is preserved, you have a tendency to express yourself more freely and vice versa.

At the European level, limits to the intrusiveness of digital evaluation on citizen's rights were posed by the G.D.P.R. The Regulation reduces the room for the freedom of companies in their management of data. Moreover, it is necessary to observe, as we have seen, the lack of a unified, communal regulatory body and the presence of different European C.N.I.L., the national Committee of informatics and liberties, who are in charge of regulating the global databanks.

It should be stressed that there is a discrepancy in remedies for data protection and algorithm regulation, as the former involves individual rights, being human rights, and the latter involves the risks associated with algorithms, which affect groups of people[51]. The same discrepancy concerns other aspects. The Regulatory authorities have requested that the reasons for collecting and processing data must be justified *a priori*, but deep learning does not have this same type of regulation, meaning its reasons for being are left to be unknown. It should be added that A.I. finds unexpected correlations between data that would otherwise appear disconnected and irrelevant from one another. Any restriction on data collection is detrimental to who runs these types of programs.

The G.D.P.R. is expected to widen the legislative gap between Europe and the free reign of online web giants from America and China, especially seeing as there is no European-origin digital giant. Paradoxically, strict

legislation on competition and privacy protection leads to European digital subordination. On the one hand, Europeans have the most protective legislation in the world for consumer protection and privacy, but, on the other hand, we are becoming a digital colony of the American and Chinese A.I. industries.

The ongoing legislation battle is over the proper level of privacy and technological freedom, and the European error would be to focus only on consumer protection, contrary to what the United States and China do, which protect large digital industries. If done, Europe would thus suffocate operators stifling any emergence of meaningful and relevant European technological startups, seeing as there are still no European *unicorns*: technology startups valued at at least one billion dollars, as rare as the mythological creatures for which they are named[52].

However, according to the principles and provisions of G.D.P.R., there are some criticisms concerning personal data protection.

First, the problem of the purpose of the technology is present, seeing as, in Europe, A.I. functions can only be carried out for specific purposes. With respect to this aspect, it is necessary to warn against the dangers posed by A.I. which, in the early stages of its operation, processes the data for predetermined purposes, but that, in the field of machine learning and its ability to adapt and consequently change its behaviours, it could begin to process those same data for purposes other than those set out in advance; all this cannot escape the control of the data subject and Data Controller.

The second aspect to review concerns the legal basis in processing data. In addition to pre-established purposes, processing can only take place if there are adequate legal bases that make that processing lawful[53].

So, when data processing does not conform to the defined legal basis expressed in the contractual obligations between the data subject and the data controller, which happens when the A.I. system escapes proper human (or algorithmic) oversight, it is difficult to find an additional legal basis. It is important to underline Rule 22 of the G.D.P.R. which establishes the need for a human subject behind automated processes in order to protect the data subjects' rights and freedoms and legitimate interests[54].

The third aspect which comes into focus is the issue of clarity of roles. And indeed, the functioning of A.I. systems presupposes the involvement of a large number of subjects (e.g. data subjects, data controllers, providers of ancillary services, third parties to whom data are disclosed for certain purposes and who can become either data processors or even the new data controllers themselves etc.).

On closer inspection, in the A.I. sector, it often happens that the privacy roles of each subject are not well defined.

Another difficult aspect is that of the processing of multiple information by A.I. systems. Oddly enough, it is not uncommon for such systems to also obtain sensitive personal data from the processing of non-personal data, such as health or sex life attitudes, from the processing of non-personal data[55]

The last aspect of being analyzed is that of controls and audits since the G.D.P.R. foresees that appropriate audits should be carried out against those who process personal data.

It must be recognized that it is not always possible to carry out controls on the functioning and processing of personal data placed on the A.I. systems. That information is often inaccessible to the data subjects who freely give up their data, which is in breach of the rights established by the G.D.P.R. which entitles the data subjects to receive information about the personal data processed or the transfer of such data to third parties and so on[56].

These are the main problems related to A.I. and data protection that if remain unaddressed, according to the respect of ethical and normative canons determined by the G.D.P.R. and other sources, can turn into unfair, advantageous opportunities for the operator[57].

Regarding the G.D.P.R., recital n. 75 speaks to the risks to the rights and freedoms of

physical persons (agents) that may result from the processing of personal data at the discretion of out-of-control A.I. systems or in the wrong hands and, therefore, liable to cause physical, material or immaterial damage. In particular, the recital warns against data processing, which may involve discrimination, theft or misuse of identity,

financial loss, damage to reputation, loss of confidentiality of personal data protected by professional secrecy, unauthorized decryption of pseudonymization, or any other significant economic or social damage.

It also provides for, where data subjects risk being deprived of their rights and freedoms or prevented from exercising control over personal data related to them or where personal data would likely be processed to reveal racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, health data or data relating to sex life or criminal convictions and crimes or related security measures.

There are cases in which the data processing leads to the assessment of personal aspects, in particular by analyzing or forecasting aspects relating to professional performance, economic situation, health, preferences or personal interests, reliability or behaviour, location or travel in order to create or use personal profiles.

In the end, there are cases in which data are processed on vulnerable physical persons, such as children, or data processing involving a large amount of personal data and data subjects.

In recital 76, it is stated that the probability and the severity of the risk to the rights and freedoms of the data subject should be determined with regard to nature, the scope, the context and purpose of the processing. The risk should be considered on the basis of an objective assessment where it is to be established if the data processing operations involve a risk or a high risk.

In recital n. 77, encouragement is given to adopting a code of conduct, approved certifications, guidelines provided by the Committee or guidance given by a Data Protection Officer.

The Committee may also issue guidelines on processing operations, which it considers unlikely to pose a high risk to the rights and freedoms of physical persons and whose measures may be sufficient in such cases to ensure that these risks are addressed[58].

It is clear that while in the past the aim was to protect the private citizen from State interference and abuse of power, protection against the misuse of personal information at the present calls into question the role of individuals who often offer their own personal data voluntarily to private companies in exchange for advantages.

Internet users, in fact, make possible, willingly or otherwise, the reconstruction of their own individual profile through cookies, tracking, and consent to the sale (or sometimes fraudulent acquisition) of their own data.

Scholars have pointed out that human rights lose their meaning, in case privacy, where their use can be traded like any other commodity in exchange for money or other advantages[59]. The free sale of privacy ends up allowing totalitarian control by those who manage this information to learn about, to pilot and to guide, through statistical analysis, the personal choices of the same users in exchange for utility. This is information given to the "public web record" on the precondition of democratic participation in online life. In this way, the logic underlying human rights would be reversed, as they would be invoked to protect individual choices as an expression of freedom, ending however to be manipulated by power centres the second in which these companies were to acquire the domain over these personal pieces of information.

The concept of "inalienable" human rights, which comes into play when political authority is able to prohibit the sale or even the free transfer of human rights, is therefore called into question, despite the implicit consent of the data subject. The authority of the State regarding its role in protecting these individual freedoms, relating to its adherence in safeguarding Human Rights, either for the individual, but moreover the entire society, is put in a complex situation because the State cannot put itself in a position to limit the sale or purchase of these individual freedoms.

The State no longer has power over that of the individual is that the great power of the individual now contrasts itself against that of the ever-diminishing authority of the State, destined to succumb to the free will of the individual over his own control of his own rights. Yet, these rights, unfortunately, are now being given to centres of political and economic power (the creators of A.I.) who are able to manipulate them in order to redirect individual choices. There is no longer the Big Brother of the State that

watches us, but there is an individual increasingly eager to be supervised[60].

6. The European and international alert system on the protection of human rights and fundamental freedoms with respect to the insides of algorithms

The Committee of Ministers of European Union adopted, on 13 February 2019 at the 1337th meeting of the Ministers' Deputies, the Declaration on the manipulative capabilities of algorithmic processes.

The interest of the Committee of Ministers is in the growing threat to the right of human beings to form opinions and make decisions independently of automated systems, which come from advanced digital technologies.

The Committee affirms that attention must be paid particularly to the capacity of digital technologies to use personal data and non-personal data to identify individual vulnerabilities and thus encourages member-States to assume their responsibilities in order to address this threat by adopting a number of measures, such as: initiating informed and inclusive public debates with a focus on providing guidance to define the difference between permissible persuasion and unacceptable manipulation; taking appropriate and proportionate measures to ensure that effective legal guarantees are in place against such forms of illegitimate interference and empowering users by promoting critical digital literacy skills, specifically, public awareness on the fact that algorithmic tools are widely used for commercial purposes and for political reasons, as well as for the wills of anti- or undemocratic processes, warfare, or direct harm.

Furthermore, the Committee points out the societal role of academia in producing independent, evidence-based and interdisciplinary research and advice for decision-makers regarding the capacity of algorithmic tools to enhance or interfere with the cognitive sovereignty of individuals, stressing the need to assess the regulatory frameworks related to political communication and electoral processes to safeguard the fairness and integrity of elections. In this regard, it should be ensured that voters have access to comparable levels of information across the political spectrum and that voters are protected effectively against unfair practices and manipulation.

In the Declaration, the Committee emphasizes that technology is an ever-growing presence in our daily lives, which prompts users to disclose personal data. There is a limited understanding regarding the use of this vast quantity of data, which are used to track personal preferences for rather unclear, and sometimes illegal, purposes. Public awareness remains limited regarding the extent to which everyday devices collect and generate a vast amount of data, that are used to train machine-learning technologies to prioritize search results, to predict and shape personal preferences and sometimes, to subject individuals to behavioural experimentation. It is fundamental to take into account the serious risks for and interests of those persons that may be especially unaware of the dangers of data exploitation, in particular children and persons belonging to marginalized communities, as well as those who are especially exposed to new forms of data-driven surveillance. Increasingly, computational means make it possible to infer intimate and detailed information about individuals from readily available data. It also facilitates the micro-targeting of individuals based on profiles in ways that may profoundly affect their lives.

Data-driven technologies and systems are designed to continuously achieve optimum solutions. When operating at scale, such processes normally prioritize certain values over others, thereby shaping the contexts and environments in which individuals, users and non-users alike, process information and make their decisions. Inevitably, such reconfiguration of environments may be beneficial for some individuals and groups but detrimental to others. The effects of the targeted use of constantly expanding volumes of aggregated data on the exercise of human rights in a broader sense, significantly beyond the current notions of personal data protection and privacy, remain understudied and require important consideration.

At present machine-learning tools have the growing capacity not only to predict choices but also to influence emotions and thoughts and alter an anticipated course of action, often subliminally. Before you go to make a purchase, Alibaba already can predict what you will buy, and in this sense, you can be the beneficiary or victim of algorithms and their ability to capture information. In this

way, Cambridge Analytica used information from Facebook to capture the voting intentions of American voters during the 2016 presidential campaign. *"There's no data like more data"* is the motto coined by the founder of Cambridge Analytica.

The Committee underlines the dangers for democratic societies that emanate from the possibility to employ such capacity to manipulate and control not only economic choices but also social and political behaviours (which has only recently become apparent). In this context, particular attention should be paid to the significant power that technological advancement confers to those, both public entities and private actors, who may use such algorithmic tools without adequate democratic oversight or control.

Fine-grained, sub-conscious and personalized levels of algorithmic persuasion may have significant effects on the cognitive autonomy of individuals and their right to form opinions and make independent decisions. Such effects remain underexplored, but notwithstanding, they cannot be underestimated, because not only may they weaken the exercise and enjoyment of individual human rights, but they may lead to the corrosion of the fundamental pillars of the Council of Europe. Its central values of human rights, democracy and the rule of law are grounded on the fundamental belief in the equality and dignity of all humans as independent moral agents.

On the international scene, the O.E.C.D. has dictated five basic principles for regulating A.I. It is a document of the general agreement aimed at setting standards, signed by 36 Member States, including the world's major economies, except China, and six non-members such as Argentina, Brazil, Colombia, Costa Rica, Peru and Romania. The O.E.C.D. first stresses that A.I. must bring benefits to people and the planet, enabling inclusive growth, sustainable development and welfare. The second principle states that A.I. systems must be designed with respect for the law, human rights, democratic values and diversity, as well as including safeguards that allow human intervention. The third principle makes it clear that A.I. systems must be transparent, and there must be a clear understanding of how they work. The fourth states that they must operate in a stable and secure manner throughout their existence and that the potential risks can be assessed continuously. Finally, the last principle requires that organizations and individuals developing, distributing or operating A.I. systems are responsible for the proper functioning in line with the above-mentioned principles.

According to some criticisms, the document, ambitious and strongly desired as it were to be, would present some inconsistencies. One of the most debated issues is the accountability of algorithms, as A.I. systems are software that can learn autonomously or make decisions without human intervention. In this sense, it is often more difficult to open the black box of deep learning software to understand the ultimate reason for a decision. In this regard, the O.E.C.D. principle of transparency could be interpreted as an obligation, which countries could include in their legislation, to develop autonomous software that is always comprehensible to man. This is a hybrid solution already adopted for most of the existing applications, in the sense that it would not really be the software to decide because, in reality, it would simply propose a decision to man, who keeps the last word and takes responsibility for the choice.

The point now is to get O.E.C.D. principles translated from policy to business to put the principles into practice. For this reason, starting in autumn 2019, the O.E.C.D. website will present an observatory of good practices and solutions for companies that will be present in this sector, because a set of rules is also a guarantee for the business itself, in the name of ethics of technological innovation, where A.I. remains secular, democratic and without preconceptions.

However, according to skeptics, it would still be early for the implementation of these rules. Because the A.I. context is still nascent, this explains the current vagueness of the laws written regarding them, since it is difficult to imagine in advance and at present the different applications of A.I. It would be better to create rules when this software appears on the market and violate existing laws. Rules and sanctions should apply to those who use A.I. for illicit purposes, always bearing in mind that it is not the technology itself who should be culpable but the person who misuses it and distorts it.

It is primarily up to the governments themselves to protect citizens with appropriate laws from the pitfalls of the web

and from the power of Palo Alto through ratifying binding international treaties, enforcing sanctions against offending States, including indiscriminate use, either public or private, of surveillance policies. There is also the non-negligible role of international human rights courts, which monitor fundamental rights[61]and human rights organizations, who act as a taser for governments, helping to enforce proper respect for universal rights. Finally, there are bodies such as the Special Rapporteur on privacy and freedom of expression, independent experts appointed by the United Nations to monitor compliance with human rights standards around the world, who submit reports to the Human Rights Council; but independent national bodies such as guarantors are not negligible.

At the company level, versus the State level, it is to be welcomed the change of ethics agreed upon by the American Business Roundtable, to which the Gotha of American capitalism has pledged itself – including the big telecommunications and digital giants like Apple, which in its mission of enterprise finally shelves the principle of profit maximization, for companies now must be held accountable not only to their shareholders, but also to their stakeholders and the other elements of the company or society who are affected by the decisions of the aforementioned, being: workers, the environment, society as a whole, and consumers. It's a socialist-style valour revolution that opposes neoliberalism, an ideology that has, starting with Friedman, guided corporate decisions, and been held hostage by predatory digital capitalism. One hopes that these initiatives will also arrive in Europe and that doing so will have positive repercussions on the respect for human rights, in practical terms, as the theoretical promises have made believe.

7. The Chinese threat: the enjoyment of human rights in the face of invasive A.I.

China, the world's second-largest economy, has an ambitious A.I. strategy to become the global leader by 2030, while already being on target to outperform the U.S. in academic research in the field by this year (2019). In this sector, geopolitics plays a leading role.

The United States uses its natural products of capitalism, Wall Street and Silicon Valley, to lead the charge in A.I. advancement. This contrasts with China's approach, which is evidenced by heavy public expenditures to finance public projects which do not seem to pay regard to fundamental rights of individual privacy. Yet, despite their differences in approach, both governments have been reluctant to pass legislation or regulation on the use of any A.I. technology.

This rivalry is similar to the Cold War. The United States feels that its A.I. technology is superior to that of China, showing a certain "complacency" about its own position. The U.S.A. falsely believes that China is capable of advancing its own A.I. technology only through Silicon Valley. Yet, Kai-Fu Lee indicates that the United States is particularly susceptible to a *technological takeover* by China, a nation equipped with a population of over one billion, over 400 million online agents to collect data from, and an authoritarian government that poses on itself no limits on privacy violation.

"See far, go further" is the motto of Hikvision, Chinese company leader in the field of facial recognition, alongside Megvii, iFlytek, Zhejiang, Dahua Technology, Meiya Pico and Yitu Technology. Hikvision works specifically with drones rigged with cameras with facial recognition technology. Facial recognition is a field in which A.I. is making important steps, and the Chinese see facial recognition as a strategic advantage. This was one of the features of Google Glass presented in 2013, which turned people's faces into business cards, revealing their identity.

Invasive applications of facial recognition can result in amoral or illicit uses, as would a system capable of creating personalized advertising based on facial recognition, turning a subject's face into a spam platform. Up until now, companies have been cautious about exploiting applications of such technology, though it does exist (Facebook has "Deep Face," Amazon "rekognition, Apple "Face ID" for the iPhone).

The Chinese, among the world's greatest fans of facial recognition, have been intent on selling this technology around the world (the Uyghurs of the Muslim Xinjiang were guinea pigs in wide-ranging trials in this field). China's facial recognition technology exists in many consumer products, but this has recently been met with disapproval from certain nations that have banned their

sale, markedly the United States. Yet, China still sells these commercial products worldwide. Additionally, China sells its facial recognition technology to authoritarian governments who wish to track their own citizens. This Chinese tech is relatively inexpensive to acquire and works quite well, being employed furtively, without public detection nor uproar.

The Chinese Government seems to want to include its 1.4 billion citizens in a database, possibly available to intelligence services, to control the population against possible disturbances. But, such a database can easily be misused by the Government itself. The gigantic Chinese archive would be vulnerable to being hacked by enemies of the State, which would compromise the identities of Chinese citizens and their movements, data points that are geotagged and timestamped.

This mapping, with different modalities in relation to the different protection of privacy, already happens in the West. Smart cameras that can count the flow of passing people reveal each person's age, gender, idle time and reaction to a given context. These data are then categorized and can be searched, even in the smallest detail, such as the colour of one's clothes, one's hair, or one's shoes. A few hundred of these devices are enough to keep an entire metropolis under control, with an average cost of 20 euros per device. Therefore only modest outlays are required when compared to the immeasurable value of the service rendered. Moreover, the technology does not change between the management of urban dynamics and the collection of data for commercial or security purposes. It is the same technology that underlies smart cities, where advanced A.I. algorithms analyze data collected from millions of sensors, generating information that can make any city efficient and minimize environmental impacts. In the Netherlands, in Eindhoven, smart street lamps have been installed with cameras and microphones that can predict the outbreak of a brawl by analyzing movements of people and the level of real noise. The same happens in Spain, Barcelona, with the installation of 1500 sensors that can even help inform waste collection and the level of air quality. In Los Angeles, thanks to sensors, it has reduced the travel time in the car by 15 percent by applying different management of traffic lights. Transforming a city's life into digital information through sensors and A.I. has undoubted benefits as you can better manage traffic, waste collection, electricity and water, etc. There are 245 million security cameras active in the world, and, according to the European Investment Bank, the smart sensor market is worth $57 billion. But many of these systems, just because they are cheap, are unreliable and run the risk of feeding the temptation to use these devices' data mining capabilities for illicit uses, such as those that would allow widespread mass controls, just as is already the case in China.

A.I. has the objective to create human intelligence in a machine. It is the engineering power (electricity) of the 21st century and rests on four fundamental factors: a great mass of data, aggressive entrepreneurs, specialized scientists and favourable policy. Starting from these four factors, it is possible to determine, between China and the U.S., who will dominate A.I. globally, for which a new bipolar world order will be created, where men will coexist with A.I.

China does not have a *Silicon Valley* and was once considered only able to copy American technology (just think that in 1975, when Microsoft was founded, China was experiencing a period of intellectual regression due to the Cultural Revolution and that in 1992 only 0.2% of the Chinese population was connected to the Internet, compared to 30% in the United States). But today China has an immense amount of data, which it draws from the real world, all the information of users– daily habits, localization and so on - in this sense, it has already far exceeded the United States so that the A.I. balance hangs in favour of China although technological colonization is attributable to the United States, where it is expected that in about 15 years the A.I. will replace about 50% of jobs.

Chinese supremacy paradoxically generates a vicious circle wherein more data produces better products, which, in turn, will produce more users who, in turn, produce more data than, in turn, improve products and so on, exponentially [62]. If today, A.I. is the new electricity, Big Data is the oil that turns on the generator and China is the largest digital data producer. Its advantage is not only quantitative, since it has more data than Europe and the United States put together, but also qualitative, since it is not just the number of users at stake but also what they do online, which is thoroughly and constantly scrutinized. The internet universe pervaded the Chinese economy. But,

this was possible due to the intervention of a leading actor: the Chinese Government. This introduced the concepts of mass entrepreneurship and mass technology, to which we also add mass data surveillance.

Favourable taxation and incentives favour investments in private startups in Chinese technology, where they are generally publicly funded. If the enterprise fails, the State is ready to eat the loss in the face of having taken the risk. State incentives guide business choices, which inherently follow the government agenda. The U.S. technology market, on the other hand, wants to remain independent, having a separation between public and private. Government support also contributes, therefore, to tilting the A.I. balance to China's side.

There are ethical issues related to A.I., such as certain choices that machines will make in certain circumstances, to which both superpowers, China and the U.S. respond differently on the basis of their own scale of values. For China, ethical issues are important problems to address, but not enough to hold back technological implementation. For example, the use of A.I. in medicine (such as the ability of machines to make precise diagnoses or biobanks, access to which would be given to doctors to carry out more effective clinical research into diseases, wherein the citizen would renounce the confidentiality on his own data in favour of the common good) or in public order (think of the predictive algorithms that manage to predict in advance how and where criminal episodes would happen) could save hundreds of lives. Promoting social good is more than enough reason for the Chinese Government to stimulate technological development. It is a techno-utilitarian approach, where technological development goes hand in hand with economic development [63].

The United States and China are aware that A.I. represents a huge competitive advantage, not only in the development of autonomous weapons and defence systems, but especially in economic terms. While the former leaves room for private businesses related to the exploitation of citizens' data by focusing on growth, China conceives A.I. as a control and management tool for citizens and focuses on government applications such as the Social Credit System (S.C.S.), a rating system that is mandatory for all Chinese citizens from 2020 and onwards. It acts as a tool for assessing the reliability of citizens on the basis of their online behaviour and takes into account parameters such as credit history, buying habits, online friends and public comments on social media, the ability to fulfill their contractual obligations and so on. This is the pervasive use of A.I[64].

Americans, while open and permissive in the digital field, are not very open to the spread of pervasive innovations, Chinese are more prone to these uses. It is possible to make purchases without money or credit card by visual identification or voice recognition alone. In this sense, the Chinese are better prepared to capture and digitize their faces and voices. To regulate traffic, many Chinese cities have a myriad of sensors and cameras that store images, while in the U.S., one is less willing to accept these mass controls that restrict privacy for public benefit[65]. The same applies to Europe, which has adopted the G.D.P.R. and, in general, is more involved in monopolies, digital security and algorithmic biases. This is another aspect of how in China, the protection of privacy and human rights, in general, gives way to profit and utilitarianism. It should, however, be recognized that China, in 2017, adopted the cybersecurity law introducing sanctions for illegal data collection, which is permitted, however, if it is public in spite of the fact that it is mass and indiscriminate[66].

Conclusions

Based on the theory of "singularity," one-day machines will be smart enough to program and improve themselves until they become independent. It is the theory of the A.I. General (A.G.I.) or the artificial superintelligence, which will create thinking machines capable of performing all the intellectual tasks of human beings, and much more.

According to Ray Kurzweil's prediction, Google's inventor, the singularity will occur in 2045 and create a self-conscious A.I. that will be one billion times more powerful than all human brains. It will be possible to transfer our memory and consciousness into microprocessors, which would allow our minds to survive biological death. Computer science and neurology will merge into one science and will represent the defeat of biological death.

This scenario arises from the distinction already in invoked between "weak" A.I., limited because it performs what it was programmed for keeping within certain limits and under human control, and "strong" A.I., super-powerful and self-aware in the human sense of the word. This last kind of intelligence could get out of the hand of its creators.

The defeat of death is an obsession for transhumanist billionaires such as Elon Musk, creator of Tesla (the first company engaged in the development of the self-driving car) who is planning the colonization of Mars by 2024.

This ambition for immortality is an accelerating factor for the development of A.I. because it is crucial to defeating death. In this field, there are those who consider death to be a disease from which it is possible to heal, for which it is believed that around 2050, immortality will be achieved (that is resistant to disease and ageing) rather than victory over death. Immortality, on the other hand, would result in the possibility of digitally tracking an individual's life, thoughts and interactions, and then downloading them to hardware and A.I. software. The brain would be disconnected from the biological body and interact with the outside world without the help of the five senses. It would be digital immortality with many ethical implications [67].

Transhumanist ideology militates, therefore, in favour of the "strong" A.I. The core of transhumanist doctrine is in idiosyncrasy for human corporeality, understood as a diminution, a biological constraint of the infinite possibilities of thought. According to transhumanists, the essence of existence is not in the body but in the ability to produce, transfer and process information[68] .

Already today, Google's search function is a self-improving system, as its machine learning algorithms constantly regulate and update the results of searches carried out by users. In fact, the possibility of systems being out of control is real.

Indeed, the aim of many projects related to A.I. is that systems can operate without human control and adapt to different situations.

Now, if the system design does not set operational boundaries related to the use of the system itself, the system may be beyond the control of even its designer. One example could be flying drones, which could escape the control of those who operate them and then cause damage. This reality is so real that, in the United States, licenses to drive them are very limited[69].

A countermeasure is to set professional standards for the development and control of intelligent systems. Designers should accurately predict the operating scope of the system and provide ways to limit the risks related to a possible overshoot of the operational scope.

Intelligent systems should be able to independently monitor whether their own operation is within limits set by their designers and enter "safe mode" or proceed to a self-monitoring shutdown if those limits were to be exceeded. It is also possible to send an alert to a human supervisor as a security mechanism, in addition to the aforementioned provision, akin to the requirement of a badge of access by a State institute.

In relation to the violation of human rights, it is necessary to develop ethical principles that can be negotiated on a computational basis and used in the face of unforeseen situations, to limit regulatory violations or to deal with unforeseeable situations with a morally significant impact.

Machines do not have morality, so they must be designed according to shared ethical rules. In this regard, affective computing, a branch of information technology that aims at the transmission of human feelings to machines, can improve the relationship between man and computer, the HCI (human-computer interaction) because a system capable of perceiving the user's state of mind can better evaluate his intentions and his real will.

In the future, this development could lead us to consider autonomous intelligent systems as a new form of life (undoubtedly created by man) that is non-biological, deserving of rights, setting off a movement for recognition and self-determinatio[70]n, very similar to what in the past has been the overcoming of slavery or the increasing protection of animals.

In this way, the intelligent machines of the future could also be, like the animals of today, the target of ethical consideration, where equal rights are demanded in accordance with legal and ethical rules. According to

Laurent Alexandre, around 2080, the world will be dominated by A.I., which will tend t[71]o merge living beings and intelligence, and will force humanity to defend the perseverance of the physical body to avoid its dissolution in the virtual world.

We do not know whether these predictions are reliable, but what is certain is that A.I. must be educated and that ethical norms must be inculcated to it since the more it will be autonomous, the more it will be called to solve moral dilemmas.

A.I., only in the hands of the right people, could make the world a better place.

References:

1. Another stage in the overcoming of man by machines was marked by the victory of the tv quiz show "Jeopardy" by the IBM program called Watson, that utilized a database of 200 million pages of facts and figures, including the entirety of Wikipedia, for a total of four terabyte of memory. Later, in 2015, a new stage in the history of AI was marked by the victory of AlphaGo, an AI developed by Deep-Mind, the property of Google, against the champion Fan Hui, one of the best Go players in the world. The difference between the game of chess and the game of Go is that the first follows mathematical logic, i.e. rational, whereas the other is based in intuitive logic. It is easier to put all the combinations of the game chess into a computer, whereas it is much more difficult to artificially create the mental universe of the game Go. For this reason the victory of AlphaGo, that at that time experts did not expect 10 or 20 years before, had marked a turning point in the history of AI.

2 An embryonic form of AI was born in the 1940s when Alan Turing created the “Enigma Machine” in order to decipher coded messages by the Germans. His discovery has given him the label of “the father of computer science.”

3 The positive aspects of nanotechnology are based on the possibility of creating microscopic devices smaller than one nanometre (e.g. a sheet of paper is about 100,000 nanometres). A technology capable of operating on this scale can generate a multitude of extraordinary applications in many fields, such as medicine where nano-antibiotics could be created, for example, to eliminate the most resistant bacteria.

4 *See* Chabert Jean-Luc et al., A History of Algorithms: From the Pebble to the Microchip (Springer 2013) 2.

5 European Commission for the Efficiency of Justice (CEPEJ), *European ethical Charter on the use of the Artificial Intelligence in judicial systems and their environment*, 69. Furthermore, according to the definition given by Wikipedia, *"AI is a discipline belonging to informatics that studies theoretical foundations, methodologies and techniques that allow the design of hardware systems and software programming systems to be able to provide the computer with performance that, to a common observer, would seem to be of exclusive relevance to human intelligence".*

6 *See generally* Laurent Alexandre, La guerra delle intelligenze. Intelligenza artificiale contro intelligenza umana (EDI 2017) 10 ss.

7 *See* Jerry Kaplan, Intelligenza artificiale. Guida al futuro prossimo (Second edition, Luiss University Press 2018) 30 ss.

8 *Id* at 104-105.

9 *Id* at 119-121.

10 European Parliament resolution of 16 February 2017 with recommendations to the Commission for civil law standards on robotics (2015/2103(INL)).

11 *See e.g.,* Patrizia Fabbri, *Cos'è l'intelligenza artificiale e quali sono le applicazioni attuali e future*, ZeroUno, 2019, (March 3, 2019, 16:00), http://www.zerounoweb.it

12 *See* Kaplan, *supra* note 3, at 153-159.

13 *See* Fabbri, *supra* note 6.

14 *See* Davide Bennato, *Quale futuro per l'umano, nel trionfo dell'intelligenza artificiale*, Cultura digitale, 2019 (March 12, 2019, 10:30), https://www.agendadigitale.eu/cultura-digitale/quale-futuro-per-lumano-nel-trionfo-dellintelligenza-artificiale/

15 That is the result of a research included in a Report of the Pew Research Center titled "Artificial intelligence and the future of human beings".

16 As far as the latter aspect is concerned, the scenario is varied. It ranges from drones capable of turning into snipers to ecoshells that allow soldiers to become superhumans, not to mention AI simulated wars, cyber attacks, robot battles, biological weapons and genetically modified insects. The introduction, then, of autonomous weapons raises numerous questions, such as that of finding the person responsible in the event that a robot commits a murder. In this case the responsibility could fall on the robot, on who is in charge of the army, but also on who designed the weapon itself. There are those who have compared AI to the atomic bomb and indeed, robotics, genomics and AI can all have the destructive power of

an atomic bomb if used with aggression and warlike intentions. From this reflection, in 2015, the non-profit organization OpenAI was founded, whose objective is to democratize AI, making available to all the expensive tools necessary for the development of AI, precisely to avoid an excessive concentration of power in the hands of only a few countries or companies.

17 *See* Bennato, *supra* note 9.

18 *Id.*

19 This development model foresees a slow start due to costs and investments and a subsequent acceleration determined by the cumulative effect and by an improvement of complementary capacities.

20 *See* Fabbri, *supra* note 8.

21 *See generally* Elena Pariotti, I diritti umani: concetto, teoria, evoluzione (Cedam 2013) 183 ss.

22 *See* Stefano Rodotà, Dal soggetto alla persona (Editoriale Scientifica 2007) 42 quoted *in* Pariotti, *supra* note 15, at 184.

23 *See* Stefano Rodotà, Tecnologia e diritti (Il Mulino 2005) 26, quoted *in* Pariotti, *supra* note 15, at 184.

24 *Id.*

25 *See* Mittellstadt Brent et al, *The Ethics of Algorithms: Mapping the Debate*, 3 Big Data & Soc'y 2, 2017.

26 Art. 2(1) Universal Declaration, art. 2(1) ICCPR, art. 2(2) ICSECR, art. 1 American Convention of Human Rights, art. 14 European Convention of Human Rights.

27 Furthermore, the European Court of Human Rights, in the case concerning the use of languages in Belgian education, has maintained that the principle of equality is undermined if the differentiated treatment has no objective and reasonable justification and that the measure that operated the differentiation must pursue a legitimate purpose and present a ratio of reasonable proportionality between the means employed and the objective pursued, and "*Caso relativo a certi aspetti delle leggi sull'uso delle lingue nell'istruzione in Belgio"*, sent. del 23 luglio 1968; *see generally* Carlo Focarelli, La persona umana nel diritto internazionale (Il Mulino 2013) 224-227.

28 General Comment n. 18, 10 November 1989, par. 13.

29 *See, e.g.,* Focarelli, *supra* note 20.

30 *See also* Giribaldi Davide, *Intelligenza artificiale, tutti i pregiudizi (bias) che la rendono pericolosa*, Agenda Digitale, 2019 (March 14, 2019, 17:00), https://www.agendadigitale.eu/cultura-digitale/intelligenza-artificiale-tutti-i-pregiudizi-bias-che-la-rendono-pericolosa/.

31 The European Commission's High-Level Expert Group on Artificial Intelligence, *Draft Ethics Guidelines for Trustworthy AI*, in http://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai.

32 *«AI is human-centric: AI should be developed, deployed and used with an "Ethical purpose" (...), grounded in and reflective of fundamental rights, societal values and the ethical principles of Beneficence (do good), Non-Maleficence (do not harm), Autonomy of humans, Justice, and Explicability», ibid.* p. 13.

33 *See generally Surden* Hurry, *Artificial Intelligence and Law: An Overview*, in Georgia State University Law Review, Vol. 35, 2019, 1326-1335.

34 *Id.*

35 *See generally* Huq Aziz .Z., "*Racial Equity in Algorithmic Criminal Justice,* in *Duke Law Journal"*, 2019, 1043 ss.

36 Kehl Danielle/ Guo Priscilla/Kessler Samuel, "*Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing, Responsive Communities Initiatives,* in *Berkman Klein Center for Internet & Society, Harvard Law School*", 2017, 9.

37 *See also* Kleiberg Jon/Lakkaraju Himabindu/Leskovec Jure/ Mullainathan Sendhil, "*Human Decision and Machine Predictions,* in *The Quarterly Journal of Economics"*, 2018, 237.

38 V. Gialuz Mitja, "*Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei "risk assessment tools" tra Stati Uniti ed Europa,* in *Diritto Penale Contemporaneo",* 2019, 3-12; see also Barile Fabio, *"Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine"*, in *Diritto Penale e Uomo*, 2019, 16-19.

39 The scandal shook the world of technology in 2018, with over 87 million personal and confidential data mysteriously passing from Facebook to Cambridge Analytica, a lobbying company founded by an American billionaire who, along with Steve Bannon, used that data to influence the American presidential vote in 2016. Cambridge Analytica took Facebook data and used them to create extremely detailed online profiles which were used to interact in a realistic way with the online community in order to manipulate public opinion in what is known as "Behavioral MicroTargeting". This types of manipulation puts privacy and democracy at risk. Mark Zuckerberg, during a sworn statement to the U.S. Congress claimed not to be aware of this traffic and committed himself to improving mechanisms for protecting confidentiality for users

of the social network, an operation that is proceeding slowly according to the U.S. authorities.

40 Giribaldi, *supra* note 24.

41 "*Will your algorithm pass the test? Create AI humans can trust. Europe to lead human-centric AI: we invite the industry, research institutes and public authorities to test ethics for trustworthy AI drafted by a group of experts. The guidelines highlight the necessity for AI to respect all applicable laws, and they purpose seven key requirements for AI development. These include among others: human oversight, transparency, privacy and fairness*", https://europe.eu/!Rh69By

42 In this regard, it should be noted that Article 15 of Directive 95/46/EC prohibits decisions based solely on automated processing, whereas Article 11 provides that *«Member States shall provide that a decision based solely on automated processing, including profiling, which produces adverse legal effects or significantly affects the data subject is prohibited unless it is authorised by Union law or by the law of the Member State to which the data subject is subject and provides adequate safeguards for the rights and freedoms of the data subject, at least the right to obtain human intervention from the data subject.»*, Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regards to the processing of personal data and on the free movement of such data (hereafter "Data Protection Directive"). This provision should be read in conjunction with Articles 5 and 6 of the C.E.D.U. on the right of access to the judge. However, automated decisions cannot be based on particular categories of personal data, such as those revealing racial or ethnic origin, political opinions, religious or philosophical beliefs or trade union membership, as well as genetic, biometric or health data, or data relating to the sexual life of the person or sexual orientation, unless there are adequate measures to safeguard rights, the freedoms and legitimate interests of the data subject. This, therefore, excludes for Europe a scenario similar to that of North America, since there is a set of rules, both at Council of Europe level and at European Union level, regulating the role of AI in decision-making processes, which remain in the hands of the individual to whom AI provides valuable but limited assistance.

43 *See, e.g.,* Mantelero Alessandro, *Come regolamentare l'intelligenza artificiale: le vie possibili*, Agenda Digitale, 2019 (March 10, 2019, 17:00), https://www.agendadigitale.eu/cultura-digitale

44 Edwards Lilian & Veale Michael, *supra* , 18 ss.

45 It should be added that many new technologies have been regulated in terms of data protection based on Article 29 Data Protection Working Party, with a supranational nature.

46 The data protection reform package therefore consists of Regulation 2016/679/EU (GDPR) and Directive 2016/680/EU, lex specialis compared to the Regulation, which replace Directive 95/46/EC, respectively, considered fundamental in the field of privacy and Framework Decision 2008/977/GAI.

47 Mantelero, *supra* note 27.

48 *Id.*

49 According to art. 4(1) of the GDPR, it is defined as personal data *"any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person"*

50 *See generally* Anselmi Niccolò & Olivi Giangiacomo, *Intelligenza artificiale e privacy, i 5 punti critici di una relazione pericolosa*, Agenda Digitale, 2019 (March 14, 2019, 16:30), https://www.agendadigitale.eu/cultura-digitale

51 *See also* Edwards Lilian & Veale Michael, Slave to the Algorithm? Why a Right to an Explanation' Is Probably Not the Remedy You Are Looking For, *16 Duke Law & Technology Review 18*, 2017, 22.

52 S*ee* Alexandre, *supra* note 2, at 22-25.

53 Art. 6, GDPR.

54 Art. 22, GDPR, *"Automated individual decision-making, including profiling. 1. The data subject shall have the right not be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. 2. Paragraph 1 shall not apply if the decision: a) is necessary for entering into, or performance of, a contract between the data subject and a data controller; b) is authorised by Union of Member-States law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedom and legitimate interest; or c) is based on the data subject's explicit consent. 3. In the cases referred to in point (a) and (c) of the paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the rights to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision [...]".*

55 *Id.* art. 9.

56 *Id.* artt. 16-22

57 *See* Anselmi & Olivi, *supra* note 32.

58 For its part, recital 78 states that the protection of the rights and freedoms of natural persons with regard to the processing of personal data requires the adoption of appropriate technical and organisational measures. It also provides for the protection of the rights and freedoms of natural persons. For which *" the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default. Such measures could consist, inter alia, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features"*. Finally, recital n. 79 states that *"the protection of the rights and freedoms of data subjects as well as the responsibility and liability of controllers and processors, also in relation to the monitoring by and measures of supervisory authorities, requires a clear allocation of the responsibilities [...]".*

59 *See generally* Focarelli, *supra* note 20, at 170-172.

60 *Id.*

61 By way of example, on 13 September 2018, the European Court of Human Rights ruled that UK laws allowing mass surveillance violate the right to privacy and freedom of expression.

62 *See generally* Kai-Fu Lee, AI Super-Powers. China, Silicon Valley and the New World Order (Houghton Mifflin Harcourt 2018) 14-50.

63 *Ib.* 102.

64 Pozzi Cristina, Benvenuti nel 2050 (Egea 2019) 39.

65 Although intelligent video surveillance systems are becoming increasingly popular in the USA. In addition, they are being distributed, for example, in schools to prevent armed attacks, but also to study student behaviour and to identify in advance those who might resort to violence. It is a sort of "predictive police" on a large scale.

66 *Kai-Fu Lee, supra* note 59, 124-125.

67 Smith Laurence C., The world in 2050 (Dutton 2010).

68 *See* Balbi Gabriele, review at Mark O'Connel, Essere una macchina (Adelphi 2019).

69 *See also* Kaplan, *supra* note 3, at 204-205.