

QUEEN'S UNIVERSITY BELFAST

European Master's Programme in Human Rights and Democratisation

A.Y. 2019/2020

The interconnection between digital and cultural revolutions:
How artificial intelligence is challenging humankind

Author: Giulia Sara Cacopardo
Supervisor: Daithí Mac Síthigh

ABSTRACT

If Karl Marx had lived during the 21st Century, he would have pointed to the new technologies and the digital revolution as the “spectrum that wanders through the world” (and no longer just Europe). The 21st century is undergoing radical changes in many ways, including the proliferation of technological means now used for any task, action or work. Unfortunately, this change cannot and must not be separated from the context in which it is embedded: a world where democratic principles, by most considered as ‘the best functioning form of government’, are being challenged by the development of an increasingly elitist and unjust policy-making, which favour a small part of the population at the expense of others considered as more vulnerable. This research aims to investigate the flaws of a system, first cultural and then legislative, regarding the use of many technologies. If it is common to think that digital and cultural revolution cannot be interconnected, nor interdependent, the author aims to prove the opposite in this paper. In two different moments of the research, the author will focus on the one hand, on the gaps in a system of thought that is still proving to be deeply erroneous and, on the other hand, on the attempts made by International Human Rights Law to fill the legislative gaps, far from being able to stem this problem. The author will not think of the technological machine as a problem within itself, but rather within the whole context in which it is inserted. The aim is to prove that the very first change must be cultural and political, given the almost silent threat that technological change brings to certain human rights. Starting with social constructs, ending with laws enacted at national and supranational level, there must be special attention to this inevitable connection between technology and the historical context.

TABLE OF CONTENTS

ABSTRACT	1
LIST OF ABBREVIATIONS	4
INTRODUCTION	7
CHAPTER ONE	9
1. Artificial Intelligence between historical evolution and strand research	9
1.1 Artificial Intelligence: basic knowledge	9
1.1.1 Eliza, PARRY and some history	9
1.2 Branches	13
1.2.1 The Algorithm	13
1.2.2 Expert systems	13
1.2.3 Fuzzy logic	15
1.2.4 Machine learning	17
1.2.5 Biometric system	18
CHAPTER TWO	22
2. New Technologies and the contemporary context	22
2.1 Before now	22
2.2 Now	23
2.1.1 The crisis of Democracy	25
2.3 New Technologies	29
2.4 Race and Gender in AI	32
CHAPTER THREE	42
3. The legal framework	42
3.1 Non-discrimination law	42
3.1.1 European Union	43
3.1.2 Council of Europe	45
3.1.3 United Nations	46
3.2 Data protection law	47
3.2.1 European Union	49
3.2.2 Council of Europe	50
3.3.3 United Nations	50
3.3 Case Law	52
3.4 Solutions?	53

3.4.1 Short term solutions	54
3.4.2 Long term solutions	56
THE UNEXPECTED PANDEMIC:.....	59
New Technologies during a global emergency	59
CONCLUSION	62
BIBLIOGRAPHY	66

LIST OF ABBREVIATIONS

AFR	Automated Facial Recognition
AI	Artificial Intelligence
CCTV	Closed-Circuit television
CEDAW	The Convention on the Elimination of All forms of Discrimination Against Women
CFREU	Charter of Fundamental Rights of the European Union
CJEU	Court of Justice of the European Union
CoE	Council of Europe
COMPAS	Correctional Offender Management Profiling for Alternative Sanctions
DPA	Data Protection Act (2018)
ECHR	European Convention on Human Rights
FRS	Facial Recognition System
GA	General Assembly
GDPR	General Data Protection Regulation
GPS	General Problem Solver
ICCPR	International Convention on Civil and Political Rights
ICERD	International Convention on the Elimination of All Forms of Racial Discrimination
ICESCR	International Covenant on Economic, Social and Cultural Rights
MS	Member States
SV	Silicon Valley
SWP	South Wales Police

UDHR

Universal Declaration of Human Rights

UN

United Nations

INTRODUCTION

Human being and machine, evolution and creation, nature and artifice: all dualisms seemingly irreconcilable that in recent decades have been symbiotically finding their own balance. Increasingly frequent are the interactions between human beings and intelligent artificial systems, which are now inextricably linked to our daily life, showing us the fastest way, recommending us the next song to listen and reminding us of the late afternoon appointment and even how to arrive to our car. Artificial Intelligence Systems are countless nowadays and, probably, while scrolling through the pages of this work, already new and equally fascinating ones will be born.

The present writing aims at – in a complex and fascinating time – giving voice to a system that is pervading our lives without any awareness of the scope of this revolution. The decade between the 1940s and the 1950s is one of the most revolutionary in the field of Artificial Intelligence (AI), and the development of technology. Alan Turing, a prominent figure in that decade, argues that “within about 50 years or so, it will be possible to schedule computers [...], to make them play the game of imitation so well that an average examiner will have no more than a 70 percent chance of accomplishing exact identification after 5 minutes of interrogation¹”. Philosophers and scientists still had several doubts about the real value assigned to the test: yet, what matters to us is to highlight that as early as 1950, when technological innovations were still taking their first steps, Turing focused on the possibility that a machine could think and that the human mind was artificially reproducible. Turing marked a historical turning point, initially in the scientific field and later, as we shall see, in something that concerns all aspects of human life such as the dissemination of digitised means of control or the emergence of machines capable of responding surprisingly quickly to human stimuli. In this research there will be a specific focus on facial recognition system, which is highly controversial and invasive. What we will try to show is that, although the technology is questionable, it depends very much on the context in which it is embedded. For this reason, in the first chapter I will analyse the main forms of new technologies such as algorithms and expert systems, so as to provide a technical basis to the reader. Through some historical hints, we will try to give a simple explanation of the main tools that guide the technologies we use daily. After having dealt with algorithms, expert systems and fuzzy logic, the last technological system we will touch upon will be the biometric one, of our particular interest as it is the basis for the development of the following chapters.

In the second chapter, starting from a brief description of the economic-governmental system that preceded the main one in force now, the ‘future’ of the democratic system will be analysed.

¹ TURING A.M., *Computing machinery and intelligence*, in *Mind*, 59, 1950, pp. 433-460.

Remembering that democracy is only one of the many forms of government that exist today, we will try to understand the inter-connection between democracy and new technologies. It will be argued, therefore, that what links all the different forms of government (in particular, China will be mentioned as a symbol of a government opposed to the democratic one and a pioneer in the technological field) is the digital revolution. A red thread that extends along all the borders - theoretical and practical. Therefore, after having a clear idea of the social system in which the new technologies are embedded, we will try to discover the weakest points of a system such as the facial recognition one. Bearing in mind the historical period we are living in - a crisis of democracy and a wave of nationalism - we will see how race and gender are two traits highly discriminated by the new technologies. If a mental, cultural and social system is unequal, racist and unfair, the new technologies that are trained by human beings can only be unequal, racist and unfair.

In the final chapter, the third, we will analyse the main legal actions that have been taken first at European level and then at supranational level, by the United Nations. Some of the limits, advantages and criticalities of the legal framework will be highlighted. By pointing out certain flaws in this system, useful alternatives to the conscious use of technologies will be proposed. Starting from these premises and giving value to the human component within the technological field, we will draw - in the conclusions – a hypothesis of virtuous use of Artificial Intelligence (AI).

The methodology used is mainly based on the analysis of literature, writings, newspaper articles and reports. Given the speed of technological development, it was necessary to use less academic and more current sources, such as newspaper articles and reports, to provide the reader with an integral view of the current situation. The literature reviewed are books, journals, reviews, statements, reports and studies. As far as the legal section is concerned, reports were mainly used, even if not exactly concerning facial recognition, they proved to be extremely useful to our research. By transposing decisions taken for other technologies, it was possible to theorize methodologies useful also for the Facial Recognition System (FRS).

CHAPTER ONE

1. Artificial Intelligence between historical evolution and strand research

1.1 Artificial Intelligence: basic knowledge

In this paragraph I am going to explore how artificial intelligence was born and what its main characteristics are. Through a brief historical explanation about the birth of the first robotic systems, we will come to understand — albeit superficially — what an algorithm, an expert system and a biometric system are. From here, the research will focus on the current applications of these systems, which bring with them the prospect of violating some of the rights recognized today, by some societies, as fundamental. Below, there is a description of some of the main technologies now in use. Although what follows may seem like a long tirade, I considered it was necessary for the reader to be made aware of some of the different types of technology in use. What will be argued is that, although there are many different types of technologies, the problem does not lie in the application of the single technology but in the unconscious and irresponsible use of all of them.

While the second chapter will focus only on the facial recognition system, this will serve as an example for all the other mechanisms. In fact, I will not limit the idea of an unhealthy use of technology only to the biometric system, but I will use it to show how the problem is structural rather than transitional. Where we are today is the result of a long process that began decades ago. I therefore believe that, in order to be able to talk about the human rights implications of using technology, it is first necessary to have a broad understating of AI.

1.1.1 Eliza, PARRY and some history

The birth of AI as a discipline has an official date recognized by the scientific community: 1956². In the summer of that year a group of scholars organized a seminar at the Dartmouth College, where the aim was to introduce the scientific community to the study of AI, a term that was first used in the seminar presentation document. This document states: “In principle, the study will proceed on the basis of conjecture since any aspect of learning or any other feature of intelligence can be described so precisely that a machine can be built to simulate them. An attempt will be made to find a way to make the machines use the language [...], solve certain types of problems and improve their

² WARWICK K. *Artificial intelligence - The basics*, Routledge, 2012.

functioning. We believe that a significant advancement can be made in one or more of these problems if a carefully selected group of scientists will work together during the summer”³. The purpose of the seminar was therefore to open a scientific debate in order to bring to light new theoretical and practical approaches, aimed at artificially reproducing the human intelligence by computers. The years following the summer of 1956 were characterized by great expectations and excitement: in 1958, some scientists created the General Problem Solver (GPS) program which, based on a software created a short time before⁴, had the ambition to imitate human problem-solving processes, extending the scope beyond purely logic-mathematical processes.

Worthy of attention is also the never realized project of J. McCarthy, one of the organizers of the Dartmouth conference⁵. In his work he presented the Advice Taker program, whose theoretical description drew what would represent a complete system of AI. This was not only able to solve logical problems, but to externalize arbitrary behaviours, to express changes to such behaviour in a simple way, and to infer some immediate consequences from whatever was communicated to him. Other contributions that are worth mentioning have been brought by less acclaimed figures but certainly of considerable importance for the Sciences.

In 1964, Joseph Weisenbaum created Eliza, the first chatterbot and one of the best English-speaking computer programs⁶. Weisenbaum put a lot of efforts into making a computer understand and converse in natural, human language: with Eliza, it was launched the first wave of artificial conversation system⁷. The main objective of Weisenbaum was to demonstrate the possibility of a conversation between a machine and a human being: Eliza was able to give substance to the scientist’s hopes by reproducing a conversation between a psychiatrist and their patient through the use of various strategies. Despite the foresight of this discovery, it was said of Eliza that “it was a computer program totally without intelligence but nonetheless remarkably good at fooling people in short conversations”⁸. However, Eliza was not the only creation of this kind. In 1972, Kenneth Colby, a psychiatrist dedicated to the theory and application of computer science and AI to psychiatry, developed a computer program called PARRY, which mimicked a person affected by schizophrenia.

³ MCCARTHY J., MINSKY M., ROCHESTER N., SHANNON C., *A proposal for the Dartmouth summer research project on artificial intelligence*, 1955, p. 1.

⁴ This is the Logic Theorist (LT) program, which is able to demonstrate logic theorems of remarkable complexity.

⁵ MCCARTHY J., *Programs with common sense*, Stanford University, Stanford, 1959.

⁶ WARWICK K., *Artificial intelligence - The basics*, Routledge, 2012, p. 2.

⁷ WARWICK K., SHAH H., *Turing’s imitation game – Conversation with the unknown*, Cambridge University Press, 2016, p. 71.

⁸ Ibid, p. 71.

With this version of the Turing test, Colby made psychiatrists compare blind PARRY utterances with those of real paranoids, and 52% of them were unable to distinguish who was who⁹.

In 1972, history of AI was about to be made: PARRY met Eliza and had the first public bot-to-bot conversation¹⁰.

Here a fragment of it:

ELIZA: Tell me your problems. Please terminate input with a period or a question mark.

PARRY: People get on my nerves sometimes.

ELIZA: I am not sure I understand you fully.

PARRY: You should pay more attention.

ELIZA: Suppose you should pay more attention.

PARRY: You're entitled to your own opinion.

ELIZA: What makes you think I am entitled to my own opinion?

PARRY: It bothers me just to be around people in general.

ELIZA: Please go on.

Overall, considerable efforts were mainly focused on the possibility of the imitation and artificial reconstruction of the human brain¹¹ and greater energies were put in trying to get the machines to speak human language rather than a mechanical language. These early discoveries have given rise to what is a very advanced field of research today. However, the aforementioned examples are not sufficient to fully explain what AI is and what its implications with the current technological mechanisms are. These considerations are the reasons why, later in years, we witness a sharp slowdown in research in the field of AI. The initial expectations began to collide with the misguided natural language translation programmes and, in addition, some theories that could work for simple examples seemed to not stick to cases more extensive and complex.

Following this initial wave of AI, during the 1970s, the field began to focus on narrower and more circumscribed aspects of AI by implementing the first expert systems. In these systems an expert human, possessing specific knowledge on a given subject, encodes some clear rules in a computer

⁹ BATACHARIA B., LEVY D., CATIZONE R., KROTOV A., WILKS Y., (1999) *CONVERSE: A Conversational Companion*, In: Wilks Y. (eds) *Machine Conversations. The Springer International Series in Engineering and Computer Science*, vol 511. Springer, Boston, MA, p. 205.

¹⁰ SHAW J., *Making Evil: The Science Behind Humanity's Dark Side*, Canongate Books Ltd, 2019.

¹¹ WARWICK K., *Artificial intelligence - The basics*, Routledge, 2012.

that can follow them to give an answer to a class of problems similar to those faced by the human expert¹². A more detailed description will follow below. With the beginning of the 1980s, it is possible to witness the birth of AI's industry: in 1982, the R1 program was designed. The R1 was the first commercial expert system, capable of designing systems and configure orders for other computers based on certain specifications. Nevertheless, progress in research did not stop with the use of AI for commercial purposes: since the mid-1980s we see the re-proposal of the so-called model based on neural networks. To talk about neural networks, it's important to go back a few decades when in the 1940s, a group of scientists, along with the birth of cybernetics¹³, began to take an interest in artificial neuron models. In 1943 the first neuronal model was created, and it was the first of its kind. The aim of studying the mechanisms of self-regulation — present in both animals and computers — was to see how both responds reactively and adaptively to external stress, changing their own behaviour. This line of research was short-lived because of a decreasing interest in cybernetics due to the ever-increasing performance of computer science. However, little by little, the research on AI passes from the stranger to the familiar and, unfailingly, begins to live with the human being. Recent innovations are not the result of the invention of new technology but the shifting forward of limits of the technologies already present on the scientific scene, from neural networks to expert systems, from self-learning programs to computing skills. Today we are facing the applications of AI in numerous research fields. AI developments, which we will briefly summarize here, are interfering with more and more strength in our daily lives and the improvement is far from over. First of all, important developments are proliferating in robotics, namely in the embodiment¹⁴ of AI systems that can experience, a real or virtual world¹⁵. In this respect, recent developments have focused on the improved robotic perceptions. In particular, more precise acoustic perception (recognition of spoken language, localization of sounds) and visual perception (recognition of objects, understanding of movements)¹⁶.

¹² BOUCHER P., *How artificial intelligence works*, Briefing of the European Parliament, March 14th, 2019, p.1.

¹³ The scientific study of how information is communicated in machines and electronic devices, comparing this with how information is communicated in the brain and nervous system,

At: <https://dictionary.cambridge.org/us/dictionary/english/device>, consulted on April 22, 2020.

¹⁴ Process by which an AI system is equipped with a physical body, so that it can relate to the outside world.

¹⁵ WARWICK K., *Artificial intelligence - The basics*, Routledge, 2012.

¹⁶ CARLUCCI AIELLO L., DAPOR M., *Intelligenza artificiale: i primi 50 anni*, in Mondo digitale, 2004, p.2.

1.2 Branches

After chronologically tracing the evolution of artificial intelligence from its beginnings up to the present day, it is necessary to discuss certain computer concepts, useful in order to understand the operation of machines and computers. This, although apparently not related to our research, turns out to be fundamental to understand the functioning of what we then investigate at a social and legal level.

1.2.1 The Algorithm

First of all, it is not possible to understand the different models of artificial intelligence – especially the ones useful to our research – without briefly discussing the concept of algorithm, fundamental for understanding how computer systems work. An algorithm is simply an abstract description of how to solve a problem: it can have inputs and can provide outputs¹⁷. To clarify what is meant by input and output, one should think of an algorithm that has to find the exit from a maze: the input data will be represented by the shape of the labyrinth, the position of the entrance and the position of the output; the output data will be represented by the suitable path that allows to overcome the labyrinth from the starting position to the final position. Between the input and the output data, there are also the so-called intermediate information (in the example of the labyrinth, the intermediate information will be the data that allow you to store during the steps of the algorithm to exit the current position it is in). Therefore, the execution of the algorithm consists in processing a set of data (input, output, intermediate information), and ‘who’ executes the algorithm is called the processor or computer¹⁸.

1.2.2 Expert systems

Having said that, we can now move forward with analysing models of Artificial Intelligence that have marked the history and evolution of this science. One of the first model that represented a turning point in research has been, as mentioned above, the creation of the so-called expert systems. The expert systems, which were named as such thanks to their ability to reason and solve problems as if they were human experts, are part of the knowledge-based systems. In these systems, knowledge is explicitly represented as a whole of statements that represent the knowledge base. This stores the

¹⁷ HAUGELAND J., *Artificial intelligence: the very idea*, MIT press, Cambridge, 1985.

¹⁸ Information on the notion of algorithm is taken from, PORTALUPI A., (ed.), *Basics of Informatics*, Zanichelli, Bologna, 2007, p. 180.

expertise of experts in the field and it is expressed in computer language that is communicated to the machine¹⁹. The structure of these systems is very simple because each rule is expressed in the basic form IF (condition) THEN (conclusion)²⁰. Obviously, it is possible that several conditions may coexist for one conclusion to be reached. An example of medical nature can help to clarify the functioning of these systems. The rule you encode can be: IF (headache) THEN (flu) or IF (headache and cough) THEN (flu)²¹, or again IF (crime) THEN (conviction). In order to reach the conclusion and thus give a solution to the problem, the expert systems use the inferential engine, a program that performs inferences from the knowledge base²².

The last element to take into consideration is the user interface, the component that makes possible the interaction between man and machine and therefore the use of the expert troubleshooting system. If we take a closer look, we realize that the functioning of an expert system involves another subject: the one using the expert system, the system user. It provides the machine with the target condition by identifying the problem that the expert system must solve and receives from the machine the description of the algorithm solving the problem²³. What are the advantages and strengths of expert systems?

Firstly, one characteristic that makes them absolutely advantageous is that they are easy to programme with the basic IF-THEN²⁴ structure code. In addition, each rule is separate from the others, requires its own data and has its own individual conclusion. This means that, if considered necessary, it is possible to add rules that the expert system will have to follow to solve the problem. This brings with it the advantage that the knowledge base is the only element to modify to update the system, adding new information, eliminating or changing the wrong ones or those that are no longer current²⁵. However, the expert system has another important advantage: it is in fact able to indicate the premises and the steps taken to reach the conclusion of the problem: it can therefore fully “justify” his behaviour. This characteristic allows the ‘backwards concatenation’²⁶, i.e. the possibility of the expert system to execute the reverse procedure: once the target has been reached, the machine is able to determine the rules and show the data and events that have occurred²⁷. In this way, going back through

¹⁹ SARTOR G., *Intelligenza artificiale e diritto. Un'introduzione*, Giuffrè, Milano, 1996, p. 16.

²⁰ WARWICK K., *Artificial intelligence - The basics*, Routledge, 2012.

²¹ Ibid.

²² SARTOR G., *Intelligenza artificiale e diritto. Un'introduzione*, cit., p. 17.

²³ SOMALVICO M., AMIGONI F., SCHIAFFONATI V., *Intelligenza artificiale*, 2019, p. 9.

²⁴ WARWICK K., *Artificial intelligence - The basics*, cit., p. 66.

²⁵ SARTOR G., *Intelligenza artificiale e diritto. Un'introduzione*, cit., p. 17.

²⁶ Different from concatenation forward, which is the normal process of the expert system whereby some events generate rules that the system follows to reach a final conclusion.

²⁷ WARWICK K., *Artificial intelligence - The basics*, cit., p. 65.

the system, it will be possible to assess whether and which new data should be entered or modified in order to achieve a specific objective. A final strength of the expert system that deserves to be highlighted is the speed of response. This gains even more weight when compared to a human expert in the same subject matter: in fact the expert system could only take a few fractions of a second to reach the resolution of the problem, whereas for a human expert it might take a few seconds or, more frequently, several minutes²⁸. However, the expert systems present some critical points such as the difficulty to collect and enter all the rules necessary for the system to reach the conclusion requested or the difficulty to bring a single standardized rule into the system. Finally, the most obvious and certain problem is that to develop a software that can solve any situation in any eventuality of the complex real world, you will need to enter a huge number of rules that would make the system too cumbersome and complicated²⁹. This problem is known as “combinatory explosion” which precisely encapsulates the need to frantically continue to add rules that can cover – no matter how highly unlikely – any eventuality that the outside world presents³⁰. The combinatorics explosion could therefore lead to a slower system, which, as said, has in the speed of resolution the greatest element that differentiates him from the human specialist.

1.2.3 Fuzzy logic

The expert systems we have analysed are based on pure logic, which means that an existing condition corresponds to an absolutely true or absolutely false fact. For this reason, such systems are useful and only make sense in confined environments that don't change over time, where the rules are almost fixed and the variables univocal³¹. However, in the real world it almost never happens like that, as Albert Einstein said: “As long as the laws of mathematics refer to reality, they are not certain, and as long as are certain, they do not refer to reality”³². In every real situation there are uncertain cases where there is no unambiguous solution but answers with a variable scope (true for a certain percentage, false for another percentage). For example, using again the medical field, a patient may have, on the basis of the data entered into the system, a certain percentage of chance of having a certain illness and a certain percentage of chance of being perfectly healthy. For these reasons the expert systems have been improved over time, in order to include situations of uncertainty and probability as well as the one of absolute certainty.

²⁸ WARWICK K., *Artificial intelligence - The basics*, cit., p. 67.

²⁹ BOUCHER P., *How artificial intelligence works, Briefing of the European Parliament*, March 14th, 2019.

³⁰ WARWICK K., *Artificial intelligence - The basics*, cit., p. 68.

³¹ BOUCHER P., *How artificial intelligence works*, cit., p. 1.

³² EINSTEIN A., *Sidelights on relativity*, 1922.

These systems are characterized by so-called fuzzy logic, according to which a certain degree of truth can be attributed to each rule. In fuzzy systems the first thing to do is to take a value from reality and blur it into a process known as fuzzification³³. So, instead of an absolute value – which hardly exist – a value that has a certain percentage between a maximum and a minimum is indicated. At this point the value is subject to the typical rules of an expert system (if-then) but the conclusion will not be absolute and always true – as it was in systems previously analysed – but will have a percentage number between a maximum and minimum value. The difference with simple expert systems can be clarified by an example: in the first case I enter the condition (if) “the water is freezing” to which the system will come to the conclusion (then) “turn on the water heater”. In a fuzzy system, the first condition is blurred into “the water is 60% icy” (which is not unbearably frosty) and the system will come to a conclusion also variable as “the water heater must be switched on and regulated according to a certain value”³⁴. In this way the fuzzy expert system can better represent models that have a closer contact with the measured reality because they allow to represent as output, solution of the results with different weights, instead of arriving at a single, incontrovertible result³⁵. The expert systems we have described – both the simple systems and the fuzzy systems – have clear application’s limits. It should be noted that all systems analysed require constant human intervention, both in the initial codification of the rules, and in possible subsequent program updates. The behaviour of the machine is determined merely by the instructions that the human specialist imparts to it and the machine learning capability is then reduced to bone³⁶. This feature reduces limits and circumscribes the degree of autonomy of these programs, which will always need human action for their development and their improvement. As a result of the above, artificial intelligence systems based on the knowledge do not seem adequate to respond to that complexity of problems where not only the variables (intercepted by fuzzy logic), but also the same rules of the game, change at all times³⁷.

The approach we have considered – that so far has had and continues to have a great success – fails to deal with situations in which human beings detect experiences from which the machine can draw to improve its awareness and “update” one’s behaviour³⁸. Indeed, this possibility is one of the fundamental characteristics of human intelligence, and the famous expression “you learn from your mistakes” highlights this aspect: our brain allows us to understand, according to the previous

³³ WARWICK K., *Artificial intelligence - The basics*, cit., p. 71.

³⁴ Ibid., p. 71.

³⁵ BONARINI A., *Sistemi Fuzzy*, in *Mondo digitale*, 1, 2003, pp. 3-14.

³⁶ In truth, even the systems we have described could be programmed in such a way as to be able to act not in accordance with the instructions received, but in more recent times new systems have been created in which self-learning is almost inherent to them.

³⁷ BOUCHER P., *How artificial intelligence works*, cit., p. 2

³⁸ WARWICK K., *Artificial intelligence - The basics*, cit., p. 135.

experiences, which functions are useful to us and which are behaviours that allow us to deal with a given situation in the best possible way. That doesn't mean that knowledge-based programs have now become obsolete and anachronistic: they remain central in some of the fields previously mentioned. Nevertheless, given the complexity of the resolution processes and decision making of the human being, who cannot be reproduced in these machines, the expert systems cannot be elevated to the designation of intelligent machines and probably - with reference to them - to the question we asked ourselves at the beginning "Can machines think?", the answer should be no.

1.2.4 Machine learning

The expert systems have therefore represented and continue to represent a key element in the development of AI. In this field, steps forward have been made and they have further revolutionized the way machines and computers operate. Before tackling these important new innovations, it is useful to introduce certain concepts that will help to better understand the nature and the functioning of new technologies. A key concept that needs to be taken into account is automatic learning (the so-called machine learning). A brief analysis of this branch of AI is crucial to the continuation of our analysis. It refers to the possibility of an algorithm to improve its performance and to learn without the human being's coding intervention³⁹: the machine will be able to learn by improving skills and functions based on past experiences. The types of 'learning machine' are different and their differentiation depends both on the algorithm used and on the purpose for which the machine was made⁴⁰. In this regard we can distinguish three large machine learning classes: supervised learning that consists of providing the machine with a series of coded information in order to create a database of notions from which the machine can draw the solution of a certain problem. This type of learning is the least complex because it is previously "man-made", therefore the machine will only have to choose the better response based on the stimuli received. Secondly, the so-called unsupervised learning when machines are entered with un-coded information. In this case it will be the machine to draw on this information without having any example of its use: the machine catalogues the information, learns its meaning, organizes it and learns the result to which they lead. In this way it is able to find the best solutions for the different problems he has to solve. Finally, we talk about learning for reinforcement —certainly the most complex— when the machine is given the necessary tools that allow it to understand the characteristics of the environment that surrounds it and, through this activity, autonomously improve their own learning and their own attitudes. To do this, the machine

³⁹ BOUCHER P., *How artificial intelligence works*, cit., p. 2.

⁴⁰ Machine learning - At: sito www.intelligenzaartificiale.it, consulted on April 26th, 2020.

is supplied with systems properly made to interact with the surroundings (sensors, cameras, GPS) in such a way that it can make the best choices to adapt to the space in which it moves independently⁴¹. Machine learning has gained new life in recent times, notoriously characterized by an enormous proliferation and spread of data: the machines have been able to learn through a continuous “training” based on the quantity of data that is entered⁴². The exponential technological development of these machines was not determined by the discovery or invention of new techniques, but by the unstoppable increase in the possibility for artificial systems to store and process data⁴³. The machine learning has found applications in different situations and through different approaches. One of the most successful applications has been the creation of the so-called artificial neural networks (artificial neural networks-ANNs), whose purpose is to attempt to artificially recreate the functioning of the human brain through the analysis of reactions of artificial neurons to electrochemical impulses⁴⁴. The objective of this mathematical/informatics model is not to recreate the exact copy of a biological brain, but to rely on its *modus operandi* in order to create an artificial network of neurons⁴⁵.

1.2.5 Biometric system

Biometric is defined as the automated recognition of individuals based on their behavioural and biological characteristics⁴⁶. The term biometrics has a double meaning: on one hand it can mean the recognition of an individual based on biological and behavioural characteristics; on the other hand, it is used in fields such as biology for statistical and mathematical methods for analysing data. Or again “(...) a concise definition of biometrics is ‘the automatic recognition of a person using distinguishing traits’; a more expansive definition of biometrics is ‘any automatically measurable, robust and distinctive physical characteristic or personal trait that can be used to identify an individual or verify

⁴¹ Machine learning – At: www.intelligenzaartificiale.it, consulted on April 26th, 2020.

⁴² The recent phenomenon of the so-called Big Data is well known, so that huge amounts of data are inserted in large databases and their processing requires more accurate and complex mechanisms.

⁴³ BOUCHER P., *How artificial intelligence works*, cit., p. 2.

⁴⁴ The neuron defined as "Functional unit of the nervous system, highly specialized cell to receive, process and transmit the information to other n. or effector cells (e.g., muscle or glandular) through electrical and chemical signals" on the Internet at www.treccani.it consulted on 12 May 2020. There are about 100 billion neurons in the human brain that are connected together forming a complex network. Each individual neuron can have up to 10,000 connections.

⁴⁵ WARWICK K., *Artificial intelligence - The basics*, cit., p.139.

⁴⁶ PATO J.N., MILLETT L. I., *Biometrics Recognition: Challenges and Opportunities*, National Research Council of the National Academies, The National academy Press, Washington D.C., 2010, pp.15-23.

the claimed identity of an individual' ”⁴⁷. Therefore, despite the various definitions that can be attributed to this concept, biometrics proves itself to be an instrument for the scientific measurement of human beings; starting with the characteristic and distinctive facial features and ending with the study of behaviour. It relies on the presumption that individuals are physically and behaviourally distinctive in a number of ways⁴⁸. The concept of biometrics sees as its conceptual basis an idea of identity in close, if not absolute, relationship with the body. What happens in a biometric process is the translation of our physical existence into codes and information. The desire to identify and systematise identities manifested itself much earlier than we think, through photography⁴⁹. Around 1840, photographic shots were used precisely for the purpose of surveillance through more refined techniques. However, photography had several limitations, especially regarding the connection between bodies and identities, for obvious reasons⁵⁰ as can be the static nature of an image. The fundamental turning point that the recognition system underwent at the time was the possibility to record movements and therefore to be able to base one's research not only on some static and immobile image but also on body movements and expressions. Anthropometry⁵¹ thus emerges as a means used by the police, through both standardized bodily measurement and sophisticated archival and retrieval systems, to identify mostly criminals or other vulnerable “categories”⁵². Even though the system of distribution of detailed information of ‘vagabonds’ or other problematic identities was never fully institutionalized, it was certainly conceptualized well before the birth of extremely advanced technological mechanisms, as they can be those that are developed together with us today⁵³. Through the union of the more “traditional” methods (such as photography) with those considered more advanced and certainly more recent, we see the formation of automated human recognition,

⁴⁷ WOODWARD J. D., HORN JR. C., GATUNE J., THOMAS A., *Biometrics: A Look at Facial Recognition*, Rand Publishing, 2003.

⁴⁸ PATO J.N., MILLETT L. I., *Biometrics Recognition: Challenges and Opportunities*, National Research Council of the National Academies, The National academy Press, Washington D.C., 2010, pp. 15-23.

⁴⁹ GATES K., *The past perfect promise of Facial Recognition Technology*, Institute of Communication Research, University of Illinois at Urbana-Champaign, 2004, p. 4.

⁵⁰ Ibid., p. 4.

⁵¹ Anthropometry is the science of obtaining systematic measurements of the human body. Anthropometry first developed in the 19th century as a method employed by physical anthropologists for the study of human variation and evolution in both living and extinct populations. In particular, such anthropometric measurements have been used historically as a means to associate racial, cultural, and psychological attributes with physical properties. Specifically, anthropomorphic measurements involve the size (e.g., height, weight, surface area, and volume), structure (e.g., sitting vs. standing height, shoulder and hip width, arm/leg length, and neck circumference), and composition (e.g., percentage of body fat, water content, and lean body mass) of humans. At: <https://biologydictionary.net/anthropometry/>: Consulted the 13th March 2020.

⁵² GATES K., *The past perfect promise of Facial Recognition Technology*, cit., p. 4.

⁵³ Ibid., p. 5.

which began with semiautomated speaker recognition systems in the 1940s. Semiautomated and fully automated fingerprint, handwriting, and facial recognition systems emerged in the 1960s, as digital computers became more widespread and capable. During this decade, the U.S. – more precisely the private company called Panoramic Research Inc. in Palo Alto (CA) – developed programs to identify faces, largely used to help the military identify enemies⁵⁴; it was in fact largely funded by the US Department of Defence. Early researches towards Facial Recognition System (FRS) (see Chapter 2 § 2.3) were part of a general effort to program computers to do what humans were incapable of doing. However, the idea was to integrate humans and machines and not to transfer human skills to a machine⁵⁵. A key characteristic of the definition of biometric is the word ‘automatic’, which sees as its pioneer Woodrow Wilson Bledsoe (cofounder of Panoramic). He is considered the father of “automated reasoning” or automatic theorem proving, an arm of early AI. Saying that the system is automatic, implies that a digital computer has been involved in the process, therefore it is possible to affirm that he developed a technique that was labelled a “hybrid man-machine system”. On one hand the role of the human operator was central for extracting facial coordinates from images, on the other hand it required a machine. The technique, which had to identify a face as a three-dimensional object with its two-dimensional image⁵⁶ can be explained - in a completely reductive way - through four main steps⁵⁷:

- Manually entering into a computer, the positions of feature points in an image – feature extraction.
- A human operator would use a rand tablet to extract the coordinates of features (such as the corners of the eyes and mouth)
- The name of the person in an image was stored in a database along with facial coordinated
- Records were organized based on those measurements

The flourishing period of the ‘70s, brought within itself various creations and discoveries such as the birth of fully automated systems, based on hand geometry and fingerprinting and, for example, small advances at programming computers to recognize human faces. The whole idea was to let a device

⁵⁴ GATES K., *Our biometric future: Facial Recognition Technology and the Culture of Surveillance*, New York University Press, 2011, p. 25.

⁵⁵ Ibid., p. 26.

⁵⁶ *A survey for Facial Recognition technology* - International Journal of Scientific and Research Publications, Volume 6, Issue 7, July 2016.

⁵⁷ GATES K., *Our biometric future: Facial Recognition Technology and the Culture of Surveillance*, cit.

answer the question “Who is X?”, therefore, being able to classify facial images became the subject of experimenting: some experts did so through the digitised writing of verbal files whereas others – Kanade in 1973 - using only photographs of the face and a new flexible picture analysis scheme with feedbacks, consisting of a collection of simple subroutines⁵⁸ each of which worked on a specific part of the picture⁵⁹. Kanade’s project of automatic facial recognition, developed at Kyoto University in Japan, correctly identified fifteen out of twenty people in collections of forty photographs: from this point on the performance of face recognition systems has improved significantly⁶⁰.

The 1990s saw an explosion on the development of computerisation and the programming of identification programs mostly to identify criminals who furtively crossed borders. In fact, in 2001 the MIT (Massachusetts Institute of Technology) Technology Review named biometrics as one of the “top ten emerging technologies that will change the world”⁶¹.

After the 9/11 tragedy in the United States, this technology had great popular support in order to identify those considered to be involved in the terrorist attack⁶². Authentication may be defined as “providing the right person with the right privileges the right access at the right time.” The wide and indiscriminate use that is made of this technology today is what concerns many scholars who, although they see great potential on the one hand, on the other hand can only express great concern; we will see why in the next chapter.

⁵⁸ With subroutines we mean: “a set of instructions which perform a task within a program”. At: <https://www.oxfordlearnersdictionaries.com/definition/english/subroutine>, Consulted 20th March 2020.

⁵⁹ GATES K., *Our biometric future: Facial Recognition Technology and the Culture of Surveillance*, New York University Press, 2011.

⁶⁰ *A survey for Facial Recognition technology* - International Journal of Scientific and Research Publications, Volume 6, Issue 7, July 2016.

⁶¹ J. D. WOODWARD, JR. C. HORN, J. GATUNE, A. THOMAS, *Biometrics: A Look at Facial Recognition*, Rand Publishing, 2003.

⁶² <https://www.nytimes.com/2020/01/24/business/london-police-facial-recognition.html>

CHAPTER TWO

2. New Technologies and the contemporary context

“In his 1955 short story *Franchise*, Isaac Asimov imagined how American democracy might be radically transformed by the digital age. In the story, set in 2008, Americans’ political will is exercised not by individual citizens who stand in line to vote, but by a massive supercomputer—the Multivac—that processes an ocean of public data with inscrutable algorithms to reliably predict the outcome of this messy, partisan, costly, and all-too-corruptible process⁶³.”

2.1 Before now

In imagining the transformation of democracy by the digital revolution, Asimov was not so wrong. The speed of this technological revolution has forced us to, in a very limited period of time, change our habits, our way of thinking, our needs, and our work. We are completely immersed in a different world of which we seem to have little or no detailed knowledge; and so, we have the duty to understand, deal and learn to live with it. It is clear, however, that the difficulty of comprehend in depth the technological world stems precisely from the speed with which technology has “appropriated” us and the ease with which it has made us his assiduous employees. Initially AI was only utilised to replace humans in the so-called “three D’s”, i.e. those jobs defined by the adjectives “Dull, dirty or dangerous”⁶⁴ (the agricultural or industrial sector), while in the field of cognitive faculties, the individual maintained a great competitive advantage over artificial machines. These last ones were in fact not sufficiently advanced and developed to be able to undermine humans in activities of analysis, communication and understanding in which it could act as a thinking being; characteristic which, according to Descartes’ teaching, defines the very existence of the human being. Today this is no longer the case. AI is increasingly challenging us in purely cognitive activities, and, with ever greater impetus, it is succeeding in obtaining better and better results, with the obvious prospect of taking our place in an unimaginable number of occupations and jobs⁶⁵. We find more and

⁶³ VALLOR S., *Lessons From Isaac Asimov’s Multivac*, 2017, Available at: <https://www.theatlantic.com/technology/archive/2017/05/lessons-from-the-multivac/523773/> Consulted 25th April 2020.

⁶⁴ LIN P., ABNEY K., BEKEY G., *Robot ethics: Mapping the issues for a mechanized world*, in *Artificial intelligence*, 175, 2011, p. 944.

⁶⁵ HARARI Y.N., *21 Lessons for the 21st Century* (2018), trad.it. di M. PIANI, Bompiani editore, 2018, pp. 44-45.

more sophisticated machines in a wide variety of fields: from the service sector to research, from the military sector to medicine, health and safety and environmental studies⁶⁶.

In this chapter I am going to firstly outline the crisis of the main professed form of government – democracy – which is at the centre of contemporary debate. Afterwards, I will consider how this crisis is interlinked and related to the digital revolution, that is changing permanently the structure of Western democracies (but nonetheless, the lifestyle of the whole world). After analysing the crisis, I will point out how a fundamental principle is violated by the new technologies: the principle of non-discrimination. Every day, countless minorities – amongst which women, black communities and so forth – suffer the consequences of a system that puts them in a position of disadvantage. If the ultimate aim is attempting to regulate society and strive towards equality, this path is certainly not the right one. In order to understand the direction in which humanity seems to be moving, it is necessary to explain, albeit not entirely in depth, the historical context that precedes us.

Humankind has gone throughout many revolutions - such as the industrial one – and, right now, we are experiencing the so-called technological revolution. When talking about history, the prefix ‘post’ is broadly used, alongside words that have served as Ariadne's thread in history, to describe some of the paradigms with which each era has interfaced itself. For example, it is common to hear about post-racism, a post-humanist era or again about a post-truth politics or the post-modern era in order to define this specific historical time. But what does really entail to see the reality as a ‘post’ reality? The prefix ‘post’ is often understood – if used in this context - to indicate a situation of positive development⁶⁷. However, what we call “development” does not seem to have brought radical improvements in some of the most deeply rooted problems in world society. Using prefixes as “post” seems to qualify current societal issues as solved, even though they are still very present today as unresolved.

2.2 Now

As stated by historian Y. N Harari, the 20th century saw the struggle between fascism, communism and liberalism and between those three, we ended up believing in freedom and thus in the so-called liberal narration⁶⁸. If freedom is considered to be the pillar of life, liberalism is therefore what can guarantee the respect of human rights and try to adjust all the imperfections. Between the 90s and 2000, liberalism became the ‘right page’ of history, adopted by most of the countries in the world.

⁶⁶ LIN P., ABNEY K., BEKEY G., *Robot ethics: Mapping the issues for a mechanized world*, cit., p. 944-945.

⁶⁷ VALERA, L., *Post-humanism: beyond humanism?*, Institute of Scientific and Technological Practice, Università Campus Bio-Medico di Roma, Copyright Cuadernos de Bioética, 2014.

⁶⁸ HARARI Y.N., *21 Lessons for the 21st Century* (2018), trad.it. di M. PIANI, Bompiani editore, 2018, pp. 44-45.

China was one of the few exceptions which, in its time, decided not to liberalise its policy and for this reason it was widely criticized. However, as we shall see, after the 2008 crisis – which signed a turning point in the economic process – the positions taken by China were no longer considered as wrong as everybody thought. The aftermath of the crisis was the direct or indirect destruction of multiple lives, given that solely in the USA, that year, 2.6 million people lost their jobs⁶⁹. The consequences of this loss have been both tangible and intangible: on one hand there was a global issue with rising inequalities, poor people becoming poorer and rich people becoming richer; on the other hand, a crisis of values has been unleashed and now, 12 years later, we are still trying to rebuild what was destroyed – both in terms of problems that need short-term solutions and those that will last for decades, if not centuries. Consequently, the ‘post-era’ – mentioned above - is characterized by the liquidity of values⁷⁰, ideals are no longer flying together with flags but are crumbling. In addition, and in parallel to all this, we are experiencing the “technological revolution”, a quick advancement in the use of technological devices in every aspect of our lives. The faith in rationalism, which we believed in for centuries now, is leaving us either because we are losing the rationale to understand where to stop or because iper-rationalism is our new *credo*. Either way, we are radically changing the concept of rationality. The violence of the digital, resides mainly in the fact that threatens to make all identities as similar to each other and therefore, in the final instance, non-existent⁷¹. Aware or not of this happening, the ways in which it is taking place are underhand because they are largely invisible to society. How technology acts it is not really known, we are in fact extremely conditioned by it and we rarely notice it. The task entrusted to humanity today is to reconstruct what is essential for the human being to live with, such as ideals, values and some kind of certainties we need to hang on. When faith in values is lost and inequalities seem to grow daily, people look up to figures who seem to provide a glimmer of hope and improvement in the short term. It is during these moments that the voices of the so-called “populist leaders” rise, characters who promise wealth to those who have none and even more wealth to those who already have it. However, as a phenomenon studied by historians and political scientists, populism sits on non-existent foundations. The consequence of such a phenomenon is in fact nationalism, a form of government very well known by leaders such as Bolsonaro, Trump or Salvini. As Miguel Benasayag stated, there are some moments in history where “masses refuse any kind of complex thought and abandon

⁶⁹GOLDMAN D., *Worst year for jobs since '45*,

Available at: <https://money.cnn.com/2009/01/09/news/economy/jobs_december/>: Consulted 8th June 2020.

⁷⁰ BAUMAN Z., *Liquid Modernity*, Polity Press, 2000.

⁷¹ BENASAYAG M., *La tirannia dell'algoritmo*, Vita e Pensiero, 2020.

themselves to pure pleasure”. Now, this sheer enjoyment in populism fits perfectly with the raising of the post-democratic digital world⁷².

Before digging into our major concerns – the loss of what we considered to be the pillars of democratic society – it is necessary and fundamental to know that the technological revolution, which we will be investigating, cannot be separated and read independently from the political context. The radical transformation of the technological paradigm is, with no doubts, influencing political choices in a way that we have never witnessed before.

2.1.1 The crisis of Democracy

Democracy as a system, as a paradigm of thoughts that has characterized the majority of countries in the last decades is going through a profound crisis, from which it might not be able to survive. As demonstrated by the Democracy Index 2019⁷³, civil liberties fell from 6.35 in 2008 to 5.74 in 2019⁷⁴ and it was recorded across all regions of the world. Governments are experiencing insufficient transparency, lack of accountability or dysfunctions. At the heart of the academic debate is the question of democracy: the future of what was considered to be the best way of government or, as Winston Churchill stated, “the worst form of government, except for all the others”⁷⁵. Given that of the over 195 countries that exist in the world, only 22 are recognized as fully democratic⁷⁶, we question if it still makes sense to talk about democracy as the best way of government. If it does, the following question is whether this way of government is combinable with the raise of technological power. The viability of democracy is strictly connected to the social and economic status of the particular country taken into account⁷⁷; the relationship between these two factors is interconnected and bi-directional, in fact, they determine each other. According to Shoshana Zuboff, the social and economic status we live in could be defined as the “age of surveillance capitalism”⁷⁸, definition that names her masterful work “The age of surveillance capitalism. The fight

⁷² BENASAYAG M., *La tirannia dell’algoritmo*, Vita e Pensiero, 2020.

⁷³ *A report by the Economist Intelligence Unit*, Democracy index 2019.

⁷⁴ Ibid.

⁷⁵ CHURCHILL W., Accessed on 7th August 2020: <https://winstonchurchill.org/resources/quotes/the-worst-form-of-government/>

⁷⁶ *A report by the Economist Intelligence Unit*, Democracy index 2019.

⁷⁷ CROZIER M., HUNTINGTON S.P., WATANUKI J., *The crisis of democracy. Report on the Governability of Democracies to the Trilateral Commission*, New York University Press, 1975, pp. 39-42.

⁷⁸ Surveillance capitalism is defined as “a new economic order that claims human experience as free raw material for hidden commercial practices of extraction, prediction, and sales. A parasitic economic logic in which the production of goods and services is subordinated to a new global architecture of behavioural modification (...)”.

for a human future at the new frontier of power”, in which she draws the condition where this new socio-political dimension has risen. What she argues is that capitalism is the outcome of consumers’ needs on the basis of a particular time and space, rather than something caused by social conflict⁷⁹. The “second modernity” we live in, she says, is characterized by a need for individualisation and self-affirmation, which serves as a key to the development of surveillance capitalism⁸⁰. Democracy has walked and is still walking side by side with neo-liberalism, which sees as its core value the free market. This last one has, as a necessary consequence, capitalism, accompanied with a strong perception of individualization by the human being, who is now considered to be just capital.

Zuboff, again, believes that the neoliberalist rhetoric has allowed society to submit and have to adapt to technological ‘progress’ and not vice-versa. She moves, in fact, a harsh criticism against giants like Google and Facebook, creators and accomplices in the possible loss of our individuality and democracy. In particular, Google has monetized the “behavioural surplus” data extracted from users, subsequently reinvested in the market for predictive products, raw material for surveillance⁸¹. Through the crisis of democracy, the neoliberal view was deeply attractive, mostly because it served as a mean to evade political ownership of economic choices and it promised to impose a new kind of order where disorder was feared⁸². Therefore, what we think governs us is a democratic process while what is actually governing us is the force of the market. Nevertheless, the debate about democracy is still wide open and the attempts to preserve it are still vigorous. What we should realize is that we might be entering an era in which the democratic process is going to take a turn onto something we still have to define and discover.

At last, even though most people see in neoliberalism a way out of uncertainty, it has brought more inequalities than before, where dangerous social division suggested an even more stratified and antidemocratic future⁸³. The free market is seen as the very paradigm of freedom, and democracy emerges as a synonym for capitalism; the democratic truth is to be replaced with the market’s truth, unless there is a joint effort by world leaders in order to take a step back from undemocratic practices and turn to the original values of democracy.

⁷⁹ ZUBOFF S., *The age of surveillance capitalism. The fight for a human future at the new frontier of power*, Public Affairs New York, March 2020.

⁸⁰ ZUBOFF S., *The age of surveillance capitalism. The fight for a human future at the new frontier of power*, Public Affairs New York, March 2020.

⁸¹ BUZZANO G., *The age of surveillance capitalism by Shoshana Zuboff*, Pandora Rivista, N° 1/2020.

⁸² ZUBOFF S., *The age of surveillance capitalism. The fight for a human future at the new frontier of power*, Public Affairs New York, March 2020, p. 39.

⁸³ S. ZUBOFF, “The age of Surveillance Capitalism”, PublicAffairs New York, 2019, p. 42.

2.1.2 China: a challenge to the West

In the previous paragraphs we learned what is affecting democracy, how and why. The free market has become the driving force behind human logic, although unconsciously for many. Capitalism is what indirectly conditions many of our choices and, in this logic based on the economic market, “the strongest wins”. For that, while facing the crisis of democracy, we find China to be the greatest challenge to this political and moral system. The ‘red’ country has been, for decades now, an autocracy and it aims at being even more powerful and strong than it already is. China has very different values and ideals compared to the West and we should not require, nor expect from it to become more “western”⁸⁴. As believing in different values, we often target ‘theirs’ as more authoritarian, of surveillance and totalitarian. The fundamental characteristics of Chinese values, which should be mentioned, are self-discipline and hard work, sacrificing personal freedom for national development and, finally, the importance of harmony, an orderly society. If the western society has always relied on the principle of naturalness of rights, Asia sees them as something the individual must earn. The word “harmony” is of great importance for Chinese policy and politics; the risk of being harmonized is profoundly true mostly because it is the aim of the current president of China, Xi Jinping. The idea of harmony is closely related to concepts as obedience and stability, which are core values in context like the Chinese one⁸⁵. Indeed, it should not surprise that article 1 of the “Constitution law of the People’s Republic of China” – which is the fundamental and supreme law of China⁸⁶ - states that “...The People’s Republic of China is a socialist state under the people’s democratic dictatorship led by the working class and based on the alliance of workers and peasants. (2) The socialist system is the basic system of the People’s Republic of China. Sabotage of the socialist system by any organization or individual is prohibited”⁸⁷. What emerges from this Article is that the law is a mean to control people rather than let them free⁸⁸. The ways in which control and surveillance are applied have been and still are disparate.

⁸⁴ JACQUES M., *Can the West’s democracy survive China’s rise to dominance?*, Available at: <https://www.economist.com/open-future/2018/06/14/can-the-wests-democracy-survive-chinas-rise-to-dominance>, Consulted 20th June 2020.

⁸⁵ STRITTMATTER K., *We have been harmonized. Life in China’s surveillance State*, Costum House, 2020.

⁸⁶ QUING BA C., *Chinese Constitutional law*, Available at: https://heinonline.org/HOL/Page?collection=journals&handle=hein.journals/bracj26&id=77&men_tab=srch_results, Consulted 15th July 2020.

⁸⁷ *Constitution of People Republic of China*, adopted 4th December 1982.

⁸⁸ STRITTMATTER K., *We have been harmonized. Life in China’s surveillance State*, cit.

However, one of those ways in particular deserves to be investigated: The Social Credit System. “Accelerating the construction of a social credit system is an important basis for comprehensively implementing the scientific development view and building a harmonious Socialist society, it is an important method to perfect the Socialist market economy system, accelerating and innovating social governance, and it has an important significance for strengthening the sincerity consciousness of the members of society, forging a desirable credit environment, raising the overall competitiveness of the country and stimulating the development of society and the progress of civilization⁸⁹”.

In the early 2000, China decided to move to a score society⁹⁰ and therefore they have created the so-called social credit system. The system considers a wide range of personal factors in order to rate the individuals (or a company). It addresses not only the financial creditworthiness of individuals but also their honesty or sincerity⁹¹ and the ultimate goal is a uniform social credit system based on penalty and award mechanism⁹². Based on the score, there are certain things an individual will be able to do whereas some other will be banned. For example, if the score is low, the passport might be taken away or the individual could not be able to access a job interview. Reactions to the social credit system from all over the world have been various. A study conducted by Freie Universitat, Berlin, states that about 80% of Chinese internet users take a positive view of the governmental and commercial social credit systems in their country⁹³; whereas R. Botsman thinks that the implementation of the Social Credit System is a plan to judge the trustworthiness of its 1.3 billion residents⁹⁴. What should never be forgotten is the deep gap between the western way of thinking and the East. The latter has a much more communitarian culture, focused on the development of the whole community rather than the individual. The “ultimate goal”, in fact, seems to be that of an achievement at community level, rather than individual⁹⁵. Given that our ultimate goal is the personal realization,

⁸⁹ State Council Notice concerning Issuance of the Planning Outline for the Construction of a Social Credit System (2014-2020), *Planning Outline for the Construction of a Social Credit System*, (2014-2020), Available at <https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/> Accessed on 20th June 2020.

⁹⁰ MAC SÍTHIGH D., SIEMS M., *The Chinese Social Credit System: A model for Other Countries?*, Modern law review, 2019, p. 36.

⁹¹ Ibid., p. 37.

⁹² Ibid., p. 36.

⁹³ Team of scholars at Freie Universität Berlin surveyed 2,200 citizens, “*Study: More than two thirds of Chinese take a positive view of social credit systems in their country*”, N° 198/2018 from July 23, 2018.

⁹⁴ BOTSMAN R., *Big Data meets Big brother as China moves to rate its citizens*, 21 October 2017, <https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion>, Accessed on 18th July 2020.

⁹⁵ MESSETTI G., *Nella testa del Dragone*, Mondadori, 3 Marzo 2020.

whereas China's dream is the realization of China itself, we can state that the West is much more individualistic.

In sum, the West is facing the crisis of democracy and that of values while the East is silently growing as a global power. The red thread that links these two seemingly distant realities is the rise of the third modernity, the digital one⁹⁶. The *credo* of the digital innovation quickly turned to the language of disruption and an obsession with speed, and it fits perfectly in the neoliberalism view, the crib where surveillance capitalism could raise⁹⁷.

2.3 New Technologies

It is believed for technology to be neutral, what makes it non-neutral is the social or economic system in which it is embedded⁹⁸. The relationship between technology and the system of thought in which it is ingrained is certainly reciprocal, so that a technology will function in a certain way according to that specific paradigm, and the system of thought will have to adapt to rapid technological growth. However, it is unthinkable to maintain the same system of thought through such radical changes, especially considering how fragile democracy is right now. The risk of the quick spreading of the digitalization depends on 'who uses what and how'. The problem relies on the interpretation that is made of it. At a global level, a research on the Global Expansion of AI Surveillance⁹⁹, provides data according to which China is surely the major driver for AI surveillance but nonetheless, liberal democracies are major users of AI.

Percentage of governments deploying AI surveillance systems (monitor, track, and surveil citizens to accomplish a range of policy objectives):

- o 51% of advanced democracies
- o 37% of closed autocratic states
- o 41% of electoral autocratic/competitive autocratic states

⁹⁶ ZUBOFF S., *The age of surveillance capitalism. The fight for a human future at the new frontier of power*, Public Affairs New York, March 2020, p. 42

⁹⁷ Ibid., p. 42.

⁹⁸ LANGDON W., *Do artifacts have politics?*, MIT Press on behalf of American Academy of Arts and Science, 2009.

⁹⁹ FELDSTEIN S., *The Global Expansion of AI Surveillance*, Carnegie Endowment for International Peace, 2019.

- o 41% of electoral democracies/illiberal democracies

Thus, all political contexts run the risk of unlawfully exploiting AI surveillance technology to obtain certain political objective. In the “High-Level conference - governing Game Changer – Impacts of artificial intelligence development on human rights, democracy and the rule of law”, the possible several threats that AI could infer to democracy were central in the discussion. The free flow of information and the free access we have to it, dramatically changes our notion of democracy and freedom. The way in which algorithms and the other technologies are designed creates a democratic emergency¹⁰⁰. The FRS shows well how this is unfolding. Citizens are monitored in a manner they have never experienced before and, primarily, in a non-transparent way. The FRS is used in a number of circumstances, such as the identification of ‘criminals’, at the border control, in the supermarket or, in some countries, while crossing the street. Even without going farther geographically, it was confirmed that, for example, the FRS has been used in King Cross, London, to identify pedestrians¹⁰¹. This demonstrates the danger between the proliferation of these technologies, which seem to be spreading at the uncontrollable rate.

Governments are now provided with new capacities to govern their citizens and what is surprising is the aggressive use that liberal democracies are making of this tool. Obviously, using those technologies does not mean for a country to be unlawful, but the legal paradigm in which those tools are framed will be investigated in the next chapter. In France, the port city of Marseille initiated a partnership with ZTE in 2016 to establish the Big Data of ‘Public Tranquillity project’. The goal of the program is to reduce crime by establishing a vast public surveillance network featuring an intelligence operation centre and nearly one thousand intelligent closed-circuit television (CCTV) cameras (the number will double by 2020)¹⁰². In comparison with other countries, France may be lagging behind. In the UK, according to the British Security Industry Authority, there is one CCTV camera for every 14 people¹⁰³. AI surveillance makes it likelier that democratic and authoritarian

¹⁰⁰ MCNAMEE J., Conference Report, HIGH-LEVEL CONFERENCE, *Governing the Game Changer – Impacts of artificial intelligence development on human rights, democracy and the rule of law*, conference co-organised by the Finnish Presidency of the Council of Europe Committee of Ministers and the Council of Europe, 26 – 27 February 2019, Helsinki, Finland – Finlandia Hall.

¹⁰¹ STUART F., *How Facial Recognition technology is bringing surveillance capitalism to our streets*, 2019, Available at: <https://www.opendemocracy.net/en/oureconomy/how-facial-recognition-surveillance-capitalism-streets/>

¹⁰² Ibid.

¹⁰³ Contributor CSM, *Facial Recognition Security Technology: The Facts*, 1st May 2020, <https://citysecuritymagazine.com/security-technology/facial-recognition-security-technology-the-facts/> Accessed on 19th July 2020.

governments may carry out surveillance that contravenes international human rights standards. AI and the FRS have been used, for example, by the European Union, to test a technology called “iBorderCtrl” in three countries¹⁰⁴—Greece, Hungary, and Latvia. This technology, screens migrants at border crossings. Individuals are asked questions about their countries of origin and circumstances of departure. The answers are then evaluated by an AI-based lie-detecting system. Travelers who honestly answer questions are given a code allowing them to cross. All others are transferred to human border guards for additional questioning. The technology behind iBorderCtrl is based on “affect recognition science” which purports to read facial expressions and infer emotional states in order to render legal judgments or policy decisions. Psychologists have widely criticized these tools, maintaining that it is difficult to rely on facial expressions alone to accurately determine a person’s state of mind. Despite scientific scepticism about these techniques, governments continue to explore their use¹⁰⁵. As said before, using these technologies does not necessarily mean to be unlawful or illegitimate. However, there were proof discrimination by algorithms in the FRS, which can either perform very well or, when unexpected variables are put in, fail.

Failing such a system could mean failing lives, since it is also used to detected criminality or as tool for predictive justice. FRS technology also has been unable to shake consistent gender and racial biases, which lead to elevated false positives for minorities and women— “the darker the skin, the more errors arise—up to nearly 35% for images of darker skinned women” noted Steve Lohr in the New York Times¹⁰⁶. Those biases in the AI system reflects historical pattern of discrimination¹⁰⁷, deepening and justifying historical inequalities. As a matter of fact, according to a research conducted by the ‘AI Now Institute’, recognition systems mis-categorize faces, algorithms used by predictive justice discriminate against black defendants and chatbots easily adopt racist and misogynistic language when trained on online discourse¹⁰⁸. A big problem related to this is the way in which the Big-Tech are handling such issue, which is systemic rather than confined to a single, specific area or company. Microsoft, for example, failed to take hundreds of allegations of harassment and

¹⁰⁴ THOMI L. B., *The Impact of Biometric Systems at EU Outside Borders on the Human Rights of Irregular Migrating Women*, Global Campus of Human Rights, 2018/2019.

¹⁰⁵ FELDSTEIN S., *The Global Expansion of AI Surveillance*, Carnegie Endowment for International Peace, 2019.

¹⁰⁶ Ibid.

¹⁰⁷ WEST S.M., WHITTAKER M., CRAWFORD K., *Discriminating Systems: Gender, Race and Power in AI*, AI Now Institute, 2019, Available at: <https://ainowinstitute.org/discriminatingsystems.html>, Accessed on 19th July 2020.

¹⁰⁸ VINCENT J., *Twitter taught Microsoft’s AI chatbot to be a racist asshole in less than a day*, 2016, March 24, The Verge, <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>, Accessed on 19th July 2020.

discrimination seriously, or again a lawsuit against Tesla alleged gender discrimination, retaliation and a hostile work environment¹⁰⁹.

2.4 Race and Gender in AI

Race and gender are two social structures useful to categorize the human being. The concept of race does so from a physical point of view and presumed belonging to a specific social group, while gender subdivides the individual according to the physical or social belonging of the female/male subcategory. These two traits - often and willingly subject to discrimination - are extremely interconnected and the way in which different types of discrimination (=unfair, treatment because of a person's race, sex etc.) are linked to and affect each other is called intersectionality¹¹⁰.

Outline these three concepts is useful to have a complete view of how I will deal the problem of bias within the field of artificial intelligence. In fact, gender and race are, historically, grounds for discrimination. Although the struggles are still current even though some results have been achieved, we are still facing events that make us doubt if there is an end to those battles. Moreover, if the context has changed, the subjects remain the same. Discrimination, in fact, has changed, albeit not radically, its mode and context of manifestation. Within the hundreds and thousands of technologies that surround us, we place our prejudices and fears, we reproduce on mathematical formulas how much we fear 'diversity' and how much we are still rooted in predominantly traditionalist cultures. That is to say that gender and race, in their being the creators of how any society is organized¹¹¹, filter through every issue.

Tangible proof that we are not talking about a platonic ὑπερουράνιος (hyperuranium), is offered to us by a case which took place in the United States. Eric L. Loomis - in February 2013 - was arrested in Wisconsin for resistance to a public official after being found driving a stolen car. The court, in determining the sentence to six years in prison, was assisted by the system COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), an analysis algorithm patented by an American company that, based on an interview with the person and information about their criminal

¹⁰⁹ Ibid.

¹¹⁰ 'Intersectionality' (Cambridge Dictionary, 2019)

Available at: <https://dictionary.cambridge.org/dictionary/english/intersectionality>. Accessed on 15 June 2020.

*In this section we are going to use sex labels of "male" and "female" to define gender classes since all sources use this subdivision. A more in depth research would require to specify all the subgroups related to gender issue.

¹¹¹ LORBER J., FARRELL S. A., *The social construction of gender*, SAGE Publications, 1991, p.111.

background, is capable of make an estimate of the offender's risk of recidivism¹¹². Through the inclusion of a set of objective data concerning the criminal past and conditions socio-economic and personal circumstances of the accused, as well as the submission of the person to 137 restricted questions, COMPAS measures the risk of recidivism as 'low', "medium" or "high." Loomis, once convicted, felt that his right to due process to be sentenced based on accurate information, his right to an individualized sentence proceeding and his right to due process were violated¹¹³. The case reached the Wisconsin Supreme Court which, however, in its 2016 judgment 9 confirmed the legitimacy of the use of the system COMPAS to guide the decision while cautioning against the misuse of the risk algorithmic analysis tools, which in any case must not replace the judge in his most paradigmatic prerogative, the judging function. In fact, the Court says that algorithms like the one in question do not allow to verify if the person is likely to commit a new offence: they provide instead a comparison of the defendant's information with that of similar subjects¹¹⁴. As a matter of fact, they defined the tool as a 'poor fit' for a sentencing decision¹¹⁵ and for this reason, COMPAS can be at best only one of the factors taken into account in a sentencing decision¹¹⁶. It is only by considering the algorithm as a non-exclusive factor in the decision that one can ensure that court decisions respect the right to individualised sentencing¹¹⁷. From this case which happened not long ago, it is clear how urgent it is to analyse and understand the critical points of this new technological challenge and its limits. Even though the case does not contain various elements such as race or gender, it well demonstrates how the delegation of power from man to machine in an uncontrolled manner may

¹¹² *State v. Loomis*, Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing, in Recent cases, Harv. L. Rev., 130, 2017.

COMPAS correctly predicts recidivism 61% of the time. But blacks are almost twice as likely as whites to be labelled a higher risk but not actually reoffend. It makes the opposite mistake among whites: they are much more likely than blacks to be labelled lower risk but go on to commit other crimes – Study by Prof. F. Z. BORGESIU, *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018.

¹¹³ LIU HW., LIN CF., CHEN YJ., *Beyond State v Loomis: artificial intelligence, government algorithmisation and accountability*, International Journal of Law and Information Technology, Volume 27, Issue 2, Summer 2019, p. 6.

¹¹⁴ *State v. Loomis*, 881 N.W.2d 749 (2016) the Courts states: «However, the COMPAS risk assessment does not predict the specific likelihood that an individual offender will reoffend. Instead, it provides a prediction based on a comparison of information about the individual to a similar data group».

¹¹⁵ LIU HW., LIN CF., CHEN YJ., *Beyond State v Loomis: artificial intelligence, government algorithmisation and accountability*, cit., p. 9.

¹¹⁶ *Ibid.*, p. 6.

¹¹⁷ COSTANZI C., *La matematica del processo: oltre le Colonne d'Ercole della giustizia penale*, in *Questione giustizia*, trimestrale n. 4, 2018, p. 185.

undermine human rights and the rule of law¹¹⁸. Therefore, this is a landmark case regarding the fallibility of new technologies, that are able to subjugate the individual with extreme ease. Therefore, the use of numbers and computers in the legal field - and many others – risk to affect the decision of a judge¹¹⁹.

It emerges, from this and other cases, how the use of artificial intelligence techniques in the legal field — and outside — is characterized by a lack of transparency but, despite that, there is an increasing global trend of using AI systems in court¹²⁰. The core of the problem is that due process and accountability are highly undermined¹²¹. Furthermore, technology have been used in other sectors as well (public and private). For example, public sector bodies can use AI for making decisions on eligibility for pension payments, housing assistance, unemployment benefits or predictive policing. In this latter field, algorithms are mainly used for predictive policing on a macro level incorporating strategic planning, prioritisation and forecasting; operational intelligence linking and evaluation (which may include, for instance, crime reduction activities), or again in decision-making or risk-assessments relating to individuals¹²². Obviously, from the social perspective, the usage of this systems raises some ethical concerns such as the possible inconclusive evidence leading to unjustified actions, the misguided evidence leading to bias or unfair outcomes leading to discrimination¹²³. In the private sector, AI can be used to select job applicants, and banks can use AI to decide whether to grant individual consumers credit and set interest rates for them¹²⁴, practice mostly used in China with, as discussed before, the Social Credit System.

Of course, in order to analyse how race and gender are treated in the network “society”, it is relevant to underline how they are perceived in real society, how they are constituted as a social, economic and political hierarchy¹²⁵. The concept of white supremacy, although in slightly less obvious and evident forms today, still exists and embedded in many cultures and traditions. In fact, whiteness is

¹¹⁸ LIU HW., LIN CF., CHEN YJ., *Beyond State v Loomis: artificial intelligence, government algorithmisation and accountability*, cit., pp.122–141.

¹¹⁹ Ibid., pp. 122-141.

¹²⁰ LIU HW., LIN CF., CHEN YJ., *Beyond State v Loomis: artificial intelligence, government algorithmisation and accountability*, International Journal of Law and Information Technology, Volume 27, Issue 2, Summer 2019, p. 15.

¹²¹ Ibid., p. 15.

¹²² OSWALD M., GRACE J., URWIN S., BARNES G. C., *Algorithmic risk assessment policing models: lessons from the Durham HART model and ‘Experimental’ proportionality*, Information & Communications Technology Law, 2018, 27:2, p. 3.

¹²³ Ibid., p. 10.

¹²⁴ BORGESIU F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018.

¹²⁵ NOBLE S. U., *Algorithm of oppression. How search engines reinforce racism*, New York University Press, 2018, p. 72.

still at the top of the system in almost all aspects of life. If race and hierarchy (e.g. patriarchy) are deeply rooted into the society and the cultural heritage as we sustain, the consequence is that also legislation and public policy are designed – probably unconsciously - based on hierarchical values. Both whiteness and maleness dominate not only in our real world but, as a consequence to modern technological developments, in our virtual world as well, reflecting the current climate. Evident from the search engines, for example, whoever designs the algorithms is not interested in endorsing democratic values such as the principle of non-discrimination since they are owned by privates. There is indisputably a traditionalist sentiment in one's view of the world, for which most people's thinking has seen, even though indirectly, the hierarchy of human categories. The strongest at the top and the most vulnerable at the bottom. As pointed out in the previous paragraphs, today we are experiencing an unprecedented crisis of values, in particular crisis of democratic values, which include equity, the principle of non-discrimination, the rule of law and right to privacy, among others. It seems that traditionalist sentiments and the transcription of past paradigms of thoughts (which included hierarchical and unequal values) are prevailing over principles such as non-discrimination. If it is true that we are experiencing a crisis of values, we can imagine that the digital world, which mirrors reality, absorbs and projects exactly the contemporary crisis. In 2012, before some of the relevant to this topic were published, Google – as a search engine – was found to have racist elements. In fact, when looking up for “Black Girls or women”, the most popular values were about sexualisation and pornography¹²⁶ – just now Google is trying to repair this damage. Gender and race are inevitably socially constructed and mutually constituted through science and technology. Search engine is only one of the many examples that should be brought to our attention.

In 2014, Brisha Borden was picking up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs. Meanwhile, in the summer 2013, Vernon Prater was picked up for shoplifting \$86.35 worth of tools from a nearby Home Depot store. Apparently, these two stories have nothing in common, besides the fact that they have committed a crime, even if not a major one. However, when they both booked in jail, a computer program spat out a score predicting the likelihood of each committing a future crime: Borden – who is black - was rated high risk, Prater – who is white - was rated low risk. It is worth mentioning that Prater had already been convicted for other crimes. And again, on a Thursday afternoon in January, Robert Julian-Borchak Williams was in his office at an automotive supply company when he got a call from the Detroit Police Department telling him to come to the station to be arrested. He thought

¹²⁶ NOBLE S. U., *Algorithm of oppression. How search engines reinforce racism*, New York University Press, 2018.

at first that it was a prank. When he got to the Police Department, they showed him a picture. The photo was blurry, but it was clearly not Mr. Williams. He picked up the image and held it next to his face. “No, this is not me,” Mr. Williams said. He did not know that his case may be the first known account of an American being wrongfully arrested based on a flawed match from a facial recognition algorithm, according to experts on technology and law¹²⁷. In fact, what happened is that the computer *got it wrong*; however, the computer must not allow for these errors, especially when the “it” is someone’s life. Regarding crime, FRS should not be trained to determine the fate of an individual¹²⁸, however, Predictive policing (Predpol) is expanding all around the world. It feeds immense quantities of data into an algorithm—geographic location, historic arrest levels, types of committed crimes, biometric data, social media feeds—in order to prevent crime, respond to criminal acts, or even to make predictions about future criminal activity. The police forces that relies on data models to decide where to locate support unit, ends to direct all of them in the same areas. This happens because the majorities of crimes tend to have place in poorest neighborhood, which means that more police force is going to be there. The vicious cycle that generates from this is obvious: if people are more controlled by the police, more people are going to get arrested and this feeds back into the model, creating a self-perpetuating loop of growing inequalities and algorithm-driven injustice¹²⁹. Predpol bases its predictions about crime on the geographical area, leaving aside ethnicity, social class, gender or any kind of prejudice. However, because the cities are segregated anyway, geography is a highly effective proxy for race. The result is the criminalization of poverty, where was proven that 85% of the people stop on the road involves young African American or Latino individuals¹³⁰. Another problem related to the geographical area, is linked to the so-called ‘target variable’. In order to understand what that means, we have to take a step back. A study conducted by Professor Frederik Zuiderveen Borgesius for the CoE, was able to identify six possibilities of how the application of AI can lead to discrimination. These findings are presented below¹³¹.

¹²⁷ HILL K., *Wrongfully accused by an algorithm*, 2020,

Available at: <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>, Accessed: 22 June 2020.

¹²⁸ BOULAMWINI J., GEBRU T., *Gender Shades, Intersectional Accuracy Disparities in Commercial Gender Classification*, Sorelle A. Friedler and Christo Wilson, 2018.

¹²⁹ BARTLETT J., *The people vs tech. How the internet is killing democracy*, Ebury Press, 2018.

¹³⁰ Ibid.

¹³¹ BORGESIOUS F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018, pp, 10-14.

- *Target Variables and Class Labels*

AI involves computers that find correlations in data sets¹³². It means that a computer, in order to understand – for example – which email is a spam, is going to search for some words, used as models to understand whether something is a spam or not. Therefore, the algorithm learns which related attributes or activities can serve as potential proxies for those qualities or outcomes of interest, and such outcome is called a “target variable”¹³³. While the target variable defines what data miners are looking for, ‘class labels’ divide all possible values of the target variable into mutually exclusive categories¹³⁴. If this model is transferred to actual individuals, it could be highly discriminatory in some cases

- *The training data: Labelling examples*

As already stated, the algorithm can be highly discriminatory if it is trained with discriminatory training data. Either the AI system is trained on biased data or it learns from a biased sample.

- *Data Collection*

Samples used to train the algorithm are composed of more than two-thirds by light-skinned male sample faces, therefore, the system will perform more precisely when they are identifying light-skinned people over dark-skinned people¹³⁵. Suppose, for instance, that less financially advantaged people rarely live in the city centres and must travel farther to their work than other employees. Therefore, poorer people are late for work more often than others because of traffic jams or problems with public transport. The company could choose “rarely being late often” as a class label to assess whether an employee is “good”. But if people with an immigrant background are, on average, poorer and live further from their work, that choice of class label would put people with an immigrant background at a disadvantage, even if they outperform other employees in other aspects¹³⁶.

- *Feature Selection*

This problem is related to the features that an organization selects for its AI system. Features are categories of data, which make it possible for the algorithm to make a selection and a prediction for

¹³² Ibid., pp.10-14.

¹³³ Ibid., pp.10-14

¹³⁴ BORGESIU F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018, pp. 10-14.

¹³⁵ BOULAMWINI J., GEBRU T., *Gender Shades, Intersectional Accuracy Disparities in Commercial Gender Classification*, Sorelle A. Friedler and Christo Wilson, 2018.

¹³⁶ BORGESIU F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, cit., pp. 10-14.

the uses to make the outcome more targeted¹³⁷. However, if the algorithm is biased, it will have biased targets and, thus, a biased outcome.

- *Proxies*

Algorithms could include in their encoding protected characteristics, and this happens when those characteristics are correlated with a protected class (such as skin colour or sexual orientation). Those characteristics encoded are not seen by humans. In practice, this could mean that, due to the learning outcomes of the algorithm, a member of a protected class could be excluded from a specific service, or, in the other way around, the individual could be targeted due to the protected characteristics installed¹³⁸. Unfortunately, it is very difficult to deal with this kind of indirect discrimination, because “Simply withholding these variables from the data mining exercise often removes criteria that hold demonstrable and justifiable relevance to the decision at hand”¹³⁹.

- *Intentional Discrimination*

An organization could directly discriminate, for example, based on racial origin¹⁴⁰. However, discriminate a certain group through an algorithm makes it less obvious and more difficult to catch.

What happens with algorithms used to make the FRS function, is that if the person in the photo is a white man, the software is right 99% of the time, but the darker the skin, the more errors arise — up to nearly 35% for images of darker skinned women, according to a new study that breaks fresh ground by measuring how technology works on people of different races and gender. This shows how biases from the real world are easily transmitted to machines. In 2015, for example, Google had to apologize after its image-recognition photo-app initially labelled African Americans as “gorillas”.

In 2019, algorithms from Google were included in a federal study of over 100 FRS that found they were biased, falsely identifying African American and Asian faces 10 times to 100 times more than Caucasian faces¹⁴¹. AI bias is perpetuated by data scientist who train algorithms based on patterns found in historical data: Amazon, for example, a company well-known on the global scale, had to

¹³⁷ THOMI L. B., *The Impact of Biometric Systems at EU Outside Borders on the Human Rights of Irregular Migrating Women*, Global Campus of Human Rights, 2018/2019.

¹³⁸ THOMI L. B., *The Impact of Biometric Systems at EU Outside Borders on the Human Rights of Irregular Migrating Women*, Global Campus of Human Rights, 2018/2019.

¹³⁹ BORGESIU F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018, pp. 10-14.

¹⁴⁰ Ibid.

¹⁴¹ SINGER N., METZ C., *Many Facial Recognition Systems Are Biased, Says U.S. Study*, <https://www.nytimes.com/2019/12/19/technology/facial-recognition-bias.html>, Accessed on 27th June 2020.

ditch their recruiting tool because it was discriminatory¹⁴². The Silicon Valley is today's most known place where all the technologies hence, algorithms, are created. If the 'big-tech guys' working in the SV have a certain mind-set, they will teach to their models to behave in that certain way. Presuming that the model they insert is – let's say indirectly – racist, the data will indicate that certain type of people have behaved badly. This loop generates a binary prediction for which, according to the algorithm, all people of that race will behave the same way. If AI is trained on biased data, it will reproduce those bias, and this is not limited to exceptional cases. If we think about the many jobs that will be replaced by machines (almost 77%), we can realize how dangerous it will be to eliminate human judgment. Technological design captures concept of gender and race and reproduce them¹⁴³. Darker-skinned females were the most misclassified group with an error rate of up to 34.7%. In contrast, lighter-skinned males had a maximum error rate of 0.8%.¹⁴⁴

In 2018, Joy Buolamwimi and Timnit Gebru conducted a study to find out why systems cannot correctly identify subjects with dark skin. They evaluated accuracy of the FRS examining 4 intersectional subgroups: darker females, darker males, lighter females, lighter males: the lowest accuracy was on darker females¹⁴⁵. The socially accepted view of gender and race – which is profoundly discriminatory – is then materialized through AI. Additionally, it is interesting to notice how linking gender and the FRS means that the face symbolises who you think you are. FRS, mirroring traditional and stereotypical canons with respect to gender, becomes discriminatory not only against women but also against all those who feel a 'fluid' gender. To be clearer, we fortunately live in an age that tries to break and overcome the classic male/female binary, we are trying to open our eyes and see that not everyone that has a beard necessarily feels like a man. FRS, unfortunately, is not able to have this vision less stereotyped and therefore, if you have a beard you must feel a man. Consequently, given the pervasiveness and influence that new technologies have on our thinking and actions, the result is a solidification and perpetuation of stereotypes. Another field where AI tools are being used more day by day is the predictive justice – see above *State v. Loomis*. Predictive justice, for example, can on the one hand speed up the pace of justice but, on the other hand, runs the great risk of judging individuals on the basis of biased algorithms. An algorithm, no matter how detailed it

¹⁴² VINCENT J., *Gender and Racial bias found in Amazon's facial recognition technology (again)*, Available at: <https://www.theverge.com/2019/1/25/18197137/amazon-rekognition-facial-recognition-bias-race-gender>, Consulted on 14th June 2020.

¹⁴³ COLLETT C., DILLON S., *AI and Gender. Four Proposals for Future Research*, University of Cambridge, 2019.

¹⁴⁴ Ibid.

¹⁴⁵ BOULAMWINI J., GEBRU T., *Gender Shades, Intersectional Accuracy Disparities in Commercial Gender Classification*, Sorelle A. Friedler and Christo Wilson, 2018.

may be, will hardly take into account the situation of an individual and will rather consider him or her as an abstract entity with no background.

Considering the latest events regarding discrimination – George Floyd¹⁴⁶ and the resulting protests all around the world – we are witnessing an awakening of consciousness. However, even if there is a lot of people protesting, the political elite does not miss the chance to control and monitor a certain part of the population, notably the most vulnerable part. However, even if something seems to be moving, algorithms are still showing more than once to be discriminatory both for women and black people. For example, Idemia¹⁴⁷ software, indicated that two of the latest algorithms were significantly more likely to mix up black women's faces than those of white women, or black or white men¹⁴⁸. The imperialist and normative construction of the West will therefore be built into machines. Not only are Black lives more subjected to unwarranted, rights-violating surveillance, they are also more subjected to false identification, giving the government new tools to target and misidentify individuals in connection with protest-related incidents¹⁴⁹.

A study published by MIT Media Lab, found that FRS performed worse when identifying an individual's gender if they were female or darker-skinned¹⁵⁰. In tests of three FRSs (by Microsoft, IBM and Megvii of China) led by MIT's Joy Buolamwini, showed inaccuracies in gender identification dependent on a person's skin colour¹⁵¹. Gender was misidentified in less than one percent of lighter-skinned males; in up to 7% of lighter-skinned females; up to 12% of darker-skinned

¹⁴⁶ George Floyd was (a black man) killed by police forces in Minneapolis, US, after allegedly passing a counterfeit \$20 bill.

¹⁴⁷ "A global leader software in Augmented Identity (an identity that ensures privacy and trust and guarantees secure, authenticated and verifiable transactions). They provide customized solutions in a world that evolves and advances faster than ever before. Today, information travels rapidly around the digitized space, defying boundaries and breaking the traditional codes of security". Website available at: <https://www.idemia.com/our-journey> - consulted: 30th June 2020

¹⁴⁸ SIMONITE T., *The Best Algorithms Struggle to Recognize Black Faces Equally*, Available at: <https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally/>, Consulted on 28th June 2020.

¹⁴⁹ BUOLAMWINI J., *We Must Fight Face Surveillance to Protect Black Lives*, Available at: <https://onezero.medium.com/we-must-fight-face-surveillance-to-protect-black-lives-5ffcd0b4c28a>, Consulted on 30th June 2020.

¹⁵⁰ VINCENT J., *Gender and Racial bias found in Amazon's facial recognition technology (again)*, Available at: <https://www.theverge.com/2019/1/25/18197137/amazon-rekognition-facial-recognition-bias-race-gender>, Consulted on 14th June 2020.

¹⁵¹ BUOLAMWINI J., *Facial recognition software is biased towards white men, researcher finds*, Available at: <https://www.media.mit.edu/articles/facial-recognition-software-is-biased-towards-white-men-researcher-finds/>, Consulted on 2nd July 2020.

males; and up to 35% in darker-skinner females¹⁵². Also, the FRS is not the only technology that perpetrates the gender biased view. “The same way that gendered bodies are deemed as politically important, humanoid robots are also recognised as such”. Humanoid robots signal economic prosperity and are an indication of technological expertise and development. By preserving physiological gender stereotypes in robotics, this results in an accumulative elevation of the political importance of both binary gender and AI.¹⁵³ Moreover, developers of AI are in large part male whereas those who recognize the bias in the developing are female. There is still a gender gap related to work in the technology field, and even this gap can only be perpetuated through new technologies.

In this chapter, the analysis started with the crisis that the democratic system is experiencing. If in the years of its flourish, democracy seemed to be the only existing mode for the human being, today this idea and this principle - the democratic one - are certainly questioned. The threat is twofold, coming both from the much-named digital revolution and from a cultural system that is proving to be unsuitable for the human being. Additionally, the pioneer country in the digital sector is China. This is characterised by several aspects, including the fact that it is an authoritarian system. Therefore, the technological primacy is in the hands of an authoritarian leader who, hopefully, will not set an example for other countries. Within this difficult system of thought and governance, we have analysed how two fundamental principles - non-discrimination and gender bias - are (not) included. These two principles, make clear the current situation for which it is necessary to think AI in a human rights framework. The potential and actual mistakes made by technological systems reflect, as we have already said, a problematic system of thought. In the next chapter, in fact, we will try to investigate how human rights are included in the contemporary debate on the ethical and legal regulation of new technologies.

¹⁵² BUOLAMWINI J., *Facial recognition software is biased towards white men, researcher finds*, Available at: <https://www.media.mit.edu/articles/facial-recognition-software-is-biased-towards-white-men-researcher-finds/>, Consulted on 2nd July 2020.

¹⁵³ COLLETT C., DILLON S., *AI and Gender. Four Proposals for Future Research*, University of Cambridge, 2019.

CHAPTER THREE

3. The legal framework

This chapter will be an attempt to analyse some of the existing legal framework on new technologies, with particular attention to concepts – also legally expressed – such as non-discrimination, racial and gender equality. Since it is impossible to touch upon all existing regulations, I will focus mainly on the European and United Nations model. Those systems, regardless of the context in which they are inserted, have certain shortcomings. Previously in the research, the FRS was central given the way in which it is used. In fact, we have seen that it has caused various damages and it was a threat to major values such as the principle of non-discrimination. The main issue with this technology – and by that I will include all the technologies – is that the harm exists whether it is regulated or not, and this poses an important question on multiple levels. Surely, on one hand both racial and gender discrimination are linked and depend on the concept of privacy. The right to privacy, in fact, is what should allow us to express ourselves without any kind of concern, controlling our identity and our agency. On the other hand, there are many other rights and values that are being challenged, from the right to a fair trial, to the right to be forgotten, ending with the great question of man-machine hybridisation. Besides the fact that the FRS has been used in very different ways depending on the location and the reasons why it was used, it remains true that the spectrum of use is very wide, and this is also confirmed by the fact that it exist more than one way to use this same technology (predictive police, border control, social credit system, etc.).

3.1 Non-discrimination law

The principle of non-discrimination is a milestone of international human-rights law. In fact, the international human rights legal framework has developed various mean (such as the European Convention on Human Rights and Fundamental Freedoms or the International Convention on Civil and Political Rights) in order to combat specific forms of discrimination¹⁵⁴. Non-discrimination law guarantees the equal enjoyment of rights and opportunities by every individual and, therefore, it is essential in order to enjoy all rights.

¹⁵⁴United Nations, <<https://www.un.org/ruleoflaw/thematic-areas/human-rights/equality-and-non-discrimination/>> Accessed on 15th July 2020.

3.1.1 European Union

The Court of Justice of the EU (CJEU) for over 50 years has consistently defined discrimination as the ‘application of different rules to comparable situations or the application of the same rule to different situations’¹⁵⁵. The legal concept of discrimination is still interpreted and implemented in very different ways and it could be, therefore, sometimes confusing, especially when the need to draft new rules for arising issues is urgent. Despite the confusion, non-discrimination is now recognized as a fundamental right and as a general principle of EU law, by the CJEU itself and by European Human Rights Law¹⁵⁶. Non-discrimination is vertically crossed by the concept of equality, rooted in the rule of law and constitutional neutrality principles¹⁵⁷. Therefore, the European Union law finds three different traditions of thoughts regarding the concept of “non-discrimination”: as “proportional equality”, as “substantive equality” and as “protection of special categories”¹⁵⁸. The first conception, called “proportional equality” relies on the idea that the equal treatment of all individuals requires that comparable situations must not be treated in the same way unless such treatment is “objectively justified”¹⁵⁹. This first assumption could be found already sort of problematic, given that “an objectively justified” reason to treat someone highly differently from someone else and in a discriminatory way, doesn’t seem to exist, mostly because a definition of “objective motivation” is exceedingly uncertain. Moreover, the second conception -of “substantive inequality” - recognize both positives and negatives duties associated to non-discrimination. Or again the third conception tries to resolve some ambiguities by specifying settings where special categories should not become the basis for unequal treatment¹⁶⁰. The third conception results to be problematic given that, in practice, it does not resolve the gap between direct and indirect discrimination. Direct discrimination occurs when a person is treated less favourably than another one on the ground of the protected characteristics¹⁶¹;

¹⁵⁵ DE VOS M., *The European Court of Justice and the march towards substantive equality in European Union anti-discrimination law*, International Journal of Discrimination and the Law 2020, Vol. 20(1) 62–87, p. 63.

¹⁵⁶ Ibid., p. 64.

¹⁵⁷ Ibid., p. 63.

¹⁵⁸ GOODMAN B., *Discrimination, Data Sanisation and Auditing in the European Union’s General Data Protection Regulation*, Heinoline, 2016, p. 498.

¹⁵⁹ Ibid., p. 498.

¹⁶⁰ Ibid., p. 498.

¹⁶¹ DE VOS M., *The European Court of Justice and the march towards substantive equality in European Union anti-discrimination law*, cit., p. 65.

whereas indirect discrimination is defined as something that occurs when a practice is supposed to be neutral, but it reveals that the practice is discriminating individuals due to a certain characteristic¹⁶². However, indirect discrimination seems to be justified if connected to an “objective justifiable aim”. Again, using the term objective is highly reductive and inconsistent, if not dangerous for an increase in discrimination carried out indiscriminately on non-existent grounds.

In the context of European regulation, we see how The Charter of Fundamental Rights of the European Union, Article 21 §1, states that:

1. Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited.

When using the term “any discrimination”, the range of characteristics that are protected widens and it refers both to direct and indirect discrimination. Recalling the above definition of “indirect”, the one just given proves to be extremely appropriate in the context of new technologies which, on their own, appear to be neutral but, if misused, have demonstrated to be highly discriminatory. The link between non-discrimination law and new technologies will be addressed later.

Meanwhile, still regarding the definition of indirect discrimination, it is useful to our investigation to refer to the “The Racial Equality Directive 2000/43/EC”, Article 2(b):

(b) indirect discrimination shall be taken to occur where an apparently neutral provision, criterion or practice would put persons of a racial or ethnic origin at a particular disadvantage compared with other persons, unless that provision, criterion or practice is objectively justified by a legitimate aim and the means of achieving that aim are appropriate and necessary.

This directive brings within inherent limitations, such as, for example, the reduction of discrimination to racial or ethnicity alone. Another limitation can be identified, as already underlined before, in the “unless that provision, criterion or practice is objectively justified by a legitimate aim and the means of achieving that aim are appropriate and necessary”. It is extremely dangerous to allow discrimination on apparently justifiable grounds. Obviously, those are just some of the provisions and rules issued by the European Union for what concerns the principle of non-discrimination and gender equality. For the latter, in 2006 the European Parliament issued the Directive 2006/54/EC “on the

¹⁶² BORGESIU F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018, pp. 10-14.

implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation”¹⁶³.

Although it is of fundamental importance, it also seems to bring with it some limitations given the circumscription of protection just in the field of work and not in many others.

3.1.2 Council of Europe

Under the European Convention on Human Rights and Fundamental Freedoms (ECHR), a very similar definition to the one above is given: discrimination is used to refer to “objectionable or illegal discrimination, on the basis of gender, skin colour or racial origin”¹⁶⁴. Article 14 of the European Convention on Human Rights also defends the principle of non-discrimination, saying that:

“The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status”¹⁶⁵.

The Coe decided to extend Article 14 in view of the limitation it had, such as the possibility to protect only depending on other rights protected by the ECHR. In order to promote equality between men and women, it was necessary to broaden the scope of application so that Article 14 could be used as a legal instrument against racial discrimination¹⁶⁶. In 2000, Protocol 12 was added to the ECHR. According to the Coe and thanks to this protocol, the prohibition of discrimination became a free-standing right. To the principle of non-discrimination, underlined in Article 1 §1, states that “No one shall be discriminated against by any public authority on any ground (...)”¹⁶⁷.

Therefore, thanks to this expansion, the limitations provided by article 14 are removed, and the protection is wider in scope¹⁶⁸.

¹⁶³Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast) *OJ L 204*, 26.7.2006, p. 23–36

¹⁶⁴BORGESIU F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018, pp. 10-14..

¹⁶⁵ Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, 4 November 1950) (ECHR) art. 14.

¹⁶⁶ THOMI L. B., *The Impact of Biometric Systems at EU Outside Borders on the Human Rights of Irregular Migrating Women*, Global Campus of Human Rights, 2018/2019.

¹⁶⁷ Protocol No. 12 to the convention for the protection of human rights and fundamental freedoms (adopted 4 November 2000) art 1 §2

¹⁶⁸ ‘*Handbook on European non-discrimination law 2018 edition*’ (FRA European Union Agency for Fundamental Rights, 2018) 18 available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-handbook-non-discrimination-law-2018_en.pdf, Accessed on 5 June 2019

3.1.3 United Nations

The non-discrimination principle exists not only as a qualifier on the obligations of the state to ensure enjoyment of all other human rights but also as a guarantee of equality before the law and equal protection of the law, as underlined in Article 26 of the International Convention on Civil and Political Rights (ICCPR)¹⁶⁹. Having specified the respective legal paradigm for the European Union, it is right to point out that all Member States (MS) in the EU are signatories of various treaties issued by UN. On December 10th, 1948, the General Assembly (GA) of the United Nations adopted the Universal Declaration of Human Rights (UDHR). Article 1 states that “all human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood¹⁷⁰”. Moreover, Article 2 reads that “Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. Furthermore, no distinction shall be made on the basis of the political, jurisdictional or international status of the country or territory to which a person belongs, whether it be independent, trust, non-self-governing or under any other limitation of sovereignty¹⁷¹”. These first two articles are particularly relevant in relation to the issue of discrimination given that, although perhaps in a generalised manner, they defend the equity of all individuals.

Furthermore, the Universal Declaration of Human Rights includes also the International Covenant on Civil and Political Rights (ICCPR)¹⁷² and the International Covenant on Economic, Social and Cultural Rights (ICESCR)¹⁷³: their common art. 3 ensures “the rights to equality between men and women in the enjoyment of all right”.¹⁷⁴ Here, the focus is rather on the equality between man and women. Thus, at a supranational level, there are documents which, by legally binding the signatory states, place the protection of the most vulnerable categories at the centre.

¹⁶⁹ United Nations, General Assembly, *Report on the Promotion and Protection of the right to freedom of opinion and expression*, A/73/348 (29 August 2019)

¹⁷⁰ UN, GA, Res 217 A(III), *Universal Declaration of Human Rights* (adopted 10 December 1948) (UDHR) art 1

¹⁷¹ Ibid.

¹⁷² UN, GA, Res 2200 A(XXI), *International Covenant on Civil and Political Rights* (adopted 16 December 1966) 999 UNTS 171 (ICCPR).

¹⁷³ UN GA, Res 2200 A(XXI), *International Covenant on Economic, Social and Cultural Rights*, (adopted 16 December 1966), United Nations, Treaty Series, vol. 993, 3 art 3

¹⁷⁴ United Nations Office of the High Commissioner of Human Rights, ‘*Women’s Human Rights and Gender Equality*’, 2019. Available at <https://www.ohchr.org/EN/Issues/Women/WRGS/Pages/WRGSIndex.aspx>, Accessed on 29 June 2019.

○ International Convention on the Elimination of All Forms of Racial Discrimination (ICERD)

When the UDHR came into existence, there was a broad agreement that the rights it contained should be translated into legal form as treaties, which would directly bind those States that agreed to their terms¹⁷⁵. Thus, lot of treaties have been issued by the United Nations, such as the Convention on the Elimination of Discrimination Against Women (CEDAW) and the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD). In the ICERD, Article 1, “racial discrimination shall mean any distinction, exclusion, restriction or preference based on race, colour, descent, or national or ethnic origin which has the purpose or effect of nullifying or impairing the recognition, enjoyment or exercise, on an equal footing, of human rights and fundamental freedoms in the political, economic, social, cultural or any other field of public life¹⁷⁶”. In Article 2 is specified that the Convention shall not apply to distinctions made by a State party between citizens and non-citizens¹⁷⁷, and sets out the obligations of State parties to combat discrimination. As well as the obvious requirements that the State itself, at all levels, must refrain from such acts, the Convention also requires a State to take appropriate measures against racial discrimination rooted in society, including the propagation of racist ideas advocated by groups and organizations¹⁷⁸.

3.2 Data protection law

With the rise of new technologies and the chance for governments to grow the amount of information they detain about the civil society; we witness an ever-deeper imbalance in the distribution of power in liberal democracy. If this form of government already brought with itself dynamics of social inequality and marked division, today the digital revolution that we are witnessing is bringing an even more marked division, given the amount of information that circulates and the ease to find and possess them, especially by governments.

Some essential democratic features are being challenged¹⁷⁹, one of these being the principle of self-determination of each individual. Within a society, there are various way to self-determinate yourself

¹⁷⁵ United Nations, Human Rights Office of the High Commissioner, The United Nations Human Rights Treaty System, Fact Sheet No.30, Rev 1, New York and Geneva, 2012, p. 9.

¹⁷⁶ UN, GA, International Convention on the Elimination of All Forms of Racial Discrimination, Adopted and opened for signature and ratification by General Assembly resolution 2106 (XX) of 21 December 1965, art 1.

¹⁷⁷ Ibid.

United Nations, Human Rights Office of the High Commissioner, The United Nations Human Rights Treaty System, Fact Sheet No.30, Rev 1, New York and Geneva, 2012, p. 9.

¹⁷⁸ UN, GA, International Convention on the Elimination of All Forms of Racial Discrimination Adopted and opened for signature and ratification by General Assembly resolution 2106 (XX) of 21 December 1965.

¹⁷⁹ BOSCO F., CREEMERS N., FERRARIS V., GUAGNIN D., KOOPS B-J., *Reforming European data protection law*, Law, government and technology series, springer 2011, pp. 3-33.

as an individual and, nowadays, informational self-determination appears to be essential¹⁸⁰. In fact, the 19th of October 1983, the German Federal Constitutional Court in its landmark Census Decision, understood the right to informational self-determination as “the authority of the individual to decide her/himself, on the basis of the idea of self-determination, when and within what limits information about her/his private life should be communicated to others”¹⁸¹. Germany was the first to introduce such a definition¹⁸². Control on information about ‘yourself’ is a crucial precondition to live an existence that may be said ‘self-determined’, especially in what is considered to be a democratic society¹⁸³. In this context, alongside the technological revolution, a legal revolution is happening and, the need to develop laws guaranteeing privacy and enforcing data protection is increasing, given that the individual's modes of self-determination are radically changing. When dealing with data protection, it is important to bear in mind that it is not simply about informational privacy but rather about informational autonomy¹⁸⁴.

“Privacy, as a legal right, should be conceived essentially as an instrument for fostering the specific yet changing autonomic capabilities of individuals that are, in a given society at a given time, necessary for sustaining a vivid democracy”¹⁸⁵.

Therefore, in the context of our research, it is interesting to underline the importance of a concept such as privacy which, nowadays, has been challenged; in fact, it makes wonder whether we can still talk about privacy in a narrow sense, given the fluidity and public domain that our data have – considering that the concept of privacy changes across cultures¹⁸⁶. Of course, the modes of self-determination of a human being are not and will never be merely reducible to data but, despite this,

¹⁸⁰ BOSCO F., CREEMERS N., FERRARIS V., GUAGNIN D., KOOPS B.-J., *Reforming European data protection law*, Law, government and technology series, Springer 2011, pp. 3-33.

*Self-determination throughout data does not mean that the ‘self’ is reducible to the data that produces but that the individual’s control over data and information is a precondition to live a self-determined existence.

¹⁸¹ BVerfGE 65, 1 - Volkszählung Judgment of the First Senate of 15 December 1983 on the mündliche Hearing of 18 and 19 October 1983 - 1 BvR 209, 269, 362, 420, 440, 484/83 in the proceedings über the constitutional complaints.

¹⁸² HOOGHMISTRA T., *Informational Self-Determination, Digital Health, and New Features of Data Protection*, Available at: https://edpl.lexxion.eu/data/article/14259/pdf/edpl_2019_02-007.pdf.

¹⁸³ BOSCO F., CREEMERS N., FERRARIS V., GUAGNIN D., KOOPS B.-J., *Reforming European data protection law*, cit., pp. 3-33.

¹⁸⁴ TZANOU M., *Data protection as a fundamental right next to privacy? ‘Reconstructing’ a not so new right*, p.89

¹⁸⁵ ROUVROY A., POULLET Y., *The right to informational self-determination and the value of self-development. Reassessing the importance of privacy for democracy*, in *Reinventing Data Protection?*, ed. Serge Gutwirth et al. (Dordrecht: Springer, 2009), 48.

¹⁸⁶ FISCHER-HÜBNER S., HOOFNAGLE C., KRONTIRIS I., RANNENBERG K., WAIDNER M., Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2468200.

having control over them is a necessary but not sufficient condition for an integral process of self-determination¹⁸⁷.

Given this idea of what ‘data-protection’ means, we are going to investigate some of the existing regulations at European and supranational level. We will focus, as we did before, on only two or three levels, just so that we can get an idea of some of the regulation in force.

3.2.1 European Union

As seen in previous chapters, the technological revolution is moving at such a speed that the current regulatory system is struggling to keep pace. Certainly both (law and technology) are human-driven fields, but the legal framework has proved to be weak in the face of some innovations. Nonetheless, at the European Union level, the CFREU surprisingly devotes two articles to privacy and personal data, enshrining ‘data protection’ as a fundamental right and as something that add a value to the right to privacy; whereas many other Human Rights instruments do not have data protection clause. Article 7 states that “Everyone has the right to respect for his or her private and family life, home and communications” and article 8(1) “Everyone has the right to the protection of personal data concerning him or her”¹⁸⁸. For this research, it is useful to focus mainly on Article 8, which deals with the processing of personal data and, in fact in (2) underlines that these data must be processed with consent and in a fairly way. Why it is so relevant to protect data?

Digital records of human behaviour may allow AI systems to infer on individuals’ preferences as well as on sexual orientation, age, gender, religious or political views¹⁸⁹. Therefore, any possible freedom to choose who to be will be highly compromised by machines. The General Data Protection Regulation (GDPR) is the most developed data protection law in the world to this date. The GDPR is the common name for Regulation (EC) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and the free movement of such data¹⁹⁰. The scope binds public authorities as well as individual or

¹⁸⁷ ROUVROY A., POULLET Y, *The right to informational self-determination and the value of self-development. Reassessing the importance of privacy for democracy*, in *Reinventing Data Protection?*, ed. Serge Gutwirth et al. (Dordrecht: Springer, 2009), p. 51.

¹⁸⁸ EU Charter of Fundamental Rights: Charter of Fundamental Rights of the European Union [2000] OJ C364/1, art 8.

¹⁸⁹ European Commission, High level expert group on Artificial Intelligence, set up by the European commission, Ethics guidelines for trustworthy AI, 8th April 2019, p. 17.

¹⁹⁰ ZANFIR-FORTUNA G., *The general data protection for regulation, analysis and guidance for US Higher Education institutions*, Ashleigh Imus editor, 2020.

corporate whereas in other jurisdictions the state can do what it wants but the private sector is regulated.

However, the GDPR has developed mechanisms to combat algorithmic discrimination, such as the so-called data-sanitisation¹⁹¹ (e.g. “Inaccurate data, if not corrected, should be securely erased, and a tamper-proof (digitally signed) record must kept”¹⁹²). Sanitise automated decision-making means remove data that contains information about protected categories¹⁹³. If this seem to be a giant step towards the elimination of discrimination, there is still the chance that a dataset – if contains variables with which protected categories are correlated – could lead to discrimination, especially if those variables are ‘opaque’ or ‘unintentional’.

3.2.2 Council of Europe

At CoE level, a decision was issued on 9 April 2019, authorising Member States to ratify, in the interest of the European Union, the Protocol amending the Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data¹⁹⁴. The notion of data protection under the CoE meets the European Union standards, even though the ECHR regulates only the right to respect for private and family life in Article 8.

3.3.3 United Nations

In September 2018, UN Secretary General issued a strategy on New technologies. The goal was to define how the UN will support the use of these technologies to accelerate the achievement of the 2030 Sustainable Development Agenda and to facilitate their alignment with the values enshrined in the UN Charter, the Universal Declaration of Human Rights and the norms and standards of International Laws¹⁹⁵. One of the principles and values that the UN has to respect is enshrined in the UDHR, Article 12, which underlines the right to privacy and defines it as a fundamental right¹⁹⁶. With

¹⁹¹ GOODMAN B., *Discrimination, Data Sanisation and Auditing in the European Union's General Data Protection Regulation*, Heinoline, 2016, p. 502.

¹⁹² Data Sanitization, Available at: <https://www.datasanitization.org/data-sanitization-regulations/>

¹⁹³ Ibid., p. 502.

¹⁹⁴ Council Decision (EU) 2013/32 of 9 April 2019 authorising Member States to ratify, in the interest of the European Union, the Protocol amending the Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data [2013] OJ L180/60 intro, p. 5.

¹⁹⁵ United Nations Secretary-General's strategy on New Technologies, <https://www.un.org/en/newtechnologies/> Accessed on 10th July 2020.

¹⁹⁶ UN, GA, Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217 A(III) (UDHR) art 1.

the UDHR not being legally binding, the right to privacy has been incorporated into the ICCPR, Article 17. Thus, this right is now legally binding for state parties. In Report A/73/348, the High Commissioner for Human Rights and the Human Rights Council stated that any interference with privacy must meet standard of legality, necessity and proportionality. This short introduction on legal framework was deemed necessary to give an idea of how the use of sensitive data and new technologies can and could be regulated. Dwelling only on the European and supranational (UN) level does not mean that other paradigms will not be taken into account or, rather, that those mentioned are considered as the only one worthy of discussion. Rather, we have chosen only some of the rules into force so that they could serve as an example to continue our analysis and, trying to abstract ourselves from these rules, understand what the flaws and possible consequences of a system like the existing one are.

In the next paragraph I will start from a case law that took place in England, I will try to understand what the flaws in the system are and - taking into account the above discussion- understand what could be improved. In fact, this case – as we shall see later – raises important issues about the use of FRS by the police. Therefore, we will try to imagine possible solutions to the questions that arise from analysing the case. What was found to be very interesting is that if a damage to human rights or to the individual occurs independently of the existing legal system, what can be improved in general and which theoretical-philosophical but also legal approach would be better to use?

Most likely, by the time this paper is read, it will already be considered ‘old’ and with it the regulations mentioned above. If the digital revolution is taking place at a speed never witnessed before, jurisprudence is also trying to follow it with the same speed. Sometimes successfully, sometimes less so. I think it is essential to even mention the problem of the universality of the law on new technologies. If we were to consider each country, we would realize that they all have a different legal framework and, in spite of this, the risks that new technologies or even FRS alone bring with are, after all, the same. For example, San Francisco have banned police from using the FRS technology¹⁹⁷ while China uses it to monitor residents in public places¹⁹⁸. And again, Britain’s data-protection regulator said last month that U.K. police should improve how they deploy the FRS¹⁹⁹ and,

¹⁹⁷ KARI P., *San Francisco is First US City to ban police use of Facial Recognition Tech*, Available at: <https://www.theguardian.com/us-news/2019/may/14/san-francisco-facial-recognition-police-ban>

¹⁹⁸ BUCKLEY C., MOZUR P., *How China uses High-Tech surveillance to Subdue Minorities*, Available at: <https://www.nytimes.com/2019/05/22/world/asia/china-surveillance-xinjiang.html>

¹⁹⁹ Department of Digital, Culture, Media & Sports, Available at: <https://www.gov.uk/government/news/government-to-strengthen-uk-data-protection-law>

in October, a French regulator ordered schools in southern France to cancel plans to introduce FRS²⁰⁰. Different usages and view on this technology, same outcome. It therefore raises the question whether there is a need for a single and universal regulation or, instead, relativism prevails even in this context.

3.3 Case Law

This part will — through discussing a recent legal case — try to analyse which are the major problems related to the existing legal systems. Considering that there are little legal provisions, if nothing, about FRS, we will start from some UN Resolutions about other new technologies and surveillance techniques to see what could be improved. This case raises important issues about the use of Automated Facial Recognition Technology (AFR) by police forces²⁰¹.

Mr Edward Bridges, a civil liberties campaigner alleged that he had been present and caught by cameras by the South Wales Police (SWP) deploying auto-facial recognition (AFR) software on two particular occasions, even though he was not subject to any investigation or action. Mr Bridges claimed that the use of AFR by the SWP was in breach of the right to privacy as contained in Article 8 of the European Convention of Human Rights (ECHR) and that the SWP had failed to comply with the Data Protection Act 1998 and its successor, the Data Protection Act 2018 (transposition of EU instrument GDPR). Mr Bridges also claimed that the application of the AFR was in breach of the Equality Act 2010 in that the SWP failed to consider that the use of AFR might produce results that were indirectly discriminatory²⁰². The SWP claimed that the purpose of using AFR was to identify offenders that were already on their watchlist and they would have done so by extracting biometric information from faces of members of the public captured. The results were that the court found no violation of the right to privacy, even though they determined that the initial collection of biometric data by AFR was an interference with the right to privacy under Article 8(1) (Para 60). Moreover, when the question became whether the SWP had interfered with such privacy rights “in accordance

²⁰⁰ STUPP C., *EU Plans Rules for Facial-Recognition Technology*, Available at: <https://www.wsj.com/articles/eu-plans-rules-for-facial-recognition-technology-11582219726>, Accessed on 10th July 2020.

²⁰¹ *R (Bridges) v CCSWP and SSHD*, Judgment Approved by the court for handing down *R (Bridges) v CCSWP and SSHD* Neutral Citation Number: [2019] EWHC 2341 (Admin) Case No: CO/4085/2018, 04/09/2019.

<<https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf>> Accessed on 10th July 2020.

²⁰² Clarks Legal, Case analysis, https://www.clarkslegal.com/Blog/Post/Case_Analysis_R_Bridges_v_CCSWP_and_SSHD_2019_EWHC_2341 Accessed on 12th July 2020.

with law”, they found no violation. Furthermore, the Court stated that the SWP had processed the personal data of all members of the public (157). It also found the SWP had given due regard to its Equality Act obligations²⁰³. Ms Denham, the Information Commissioner, also recently released statements indicating that in July her office was conducting an investigation – published in October 2019²⁰⁴ - into the trials undertaken by the police; because she was highly concerned about the usage of AFR²⁰⁵.

The fear that she exposed are common all over the world, not just in the United Kingdom. Starting from China and the surveillance system they used in the recent Hong Kong riots²⁰⁶; ending with California, which has recently passed a law banning state and local law enforcement from using body cameras with AFR²⁰⁷, there is a growing concern about the usage of those surveillance technologies. As mentioned in the previous chapter, many more cases than the ones declared are taking place. Therefore, a long and a short-term solution are urgently needed to avoid the perpetuation of certain discriminatory and unjust paradigms of thought. The great supranational bodies, such as the United Nations, have worked and are working in trying to develop guidelines - ethical and legal - useful for a lawful and fair use of these technologies. FRS is becoming increasingly important given the latest events that have hit the globe (Covid-19) and that seem to need constant monitoring of citizens.

3.4 Solutions?

When I said that we need long-term and short-term solutions, I refer to, that on the one hand, there is a need for a legal paradigm that puts a stop mark to the perpetuation of discrimination now and immediately. However, on the other hand, we have seen how new technologies are potentially neutral

²⁰³ *R (Bridges) v CCSWP and SSHD*, Judgment Approved by the court for handing down *R (Bridges) v CCSWP and SSHD* Neutral Citation Number: [2019] EWHC 2341 (Admin) Case No: CO/4085/2018, 04/09/2019.

<<https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf>> Accessed on 10th July 2020.

²⁰⁴ Information Commissioner’s Opinion, *The Use of live Facial recognition technology by law enforcement in public places*, Available at https://icosearch.ico.org.uk/s/search.html?query=R+%28Bridges%29+v+South+Wales+Police%2C&collection=ico-meta&profile=_default

²⁰⁵ DE VERE C., *Facial Recognition in public places*, Available at: <https://www.infolaw.co.uk/newsletter/2019/10/facial-recognition-public-spaces/>, Accessed on 13th July 2020.

²⁰⁶ OVIDE S., *The Real Dangers of Surveillance*, Available at: <https://www.nytimes.com/2020/06/12/technology/surveillance-protests-hong-kong.html>

²⁰⁷ DE VERE C., *Facial Recognition in public places*, Available at: <https://www.infolaw.co.uk/newsletter/2019/10/facial-recognition-public-spaces/>, Accessed on 13th July 2020.

and basically reflect the current climate. Therefore, change must also occur on a cultural and ideological level. If we allow new technologies to undermine the democratic principles of inclusive participation and respect, we risk shattering even the last drop of humanity that seems to have remained in the world.

3.4.1 Short term solutions

The UN have tried to protect human rights, issuing some resolutions that could help to regulate the exploitation of new technologies. In those writings some events are explicit, such as how China uses the systems of targeted surveillance in an intrusive way²⁰⁸, which is dangerous not only if used in the police field, but also if used to detect some preferences, feelings, emotions and so forth²⁰⁹. A person's face is the symbol of identity and uniqueness, hence, turning the human face into an object of measurement and categorisation by an automated process could touch the right to human dignity and the very principle of a democratic society - even without the threat of it being used as a tool for oppression by an authoritarian state²¹⁰. In Resolution A/HRC/41/35 (2019) the UN stated that the development and usage of those technologies must be consistent with International human rights obligations and must be conducted on the basis of a legal framework which must be publicly accessible, clear, precise, comprehensive and non-discriminatory²¹¹. There are, to date, no laws relating - for example - to facial recognition. There is a strong tendency to adapt old legislation to new problems. Obviously, it could be a very reasonable solution if it could block the reckless use of these technologies. To limit a very general law that does not have new technology as its specific object seems useless. Governments, ONG, and civil society itself has the duty to inform citizens in a comprehensible way for everybody. Furthermore, in this Report (Para 26, p.8) – focused on targeted surveillance – it is underlined how this system created incentives for self-censorship especially by journalists, human rights activists and so forth. Bearing this in mind, we can remember how in the case held in the U.K., explained previously, reference was made to a watchlist of people who are within the algorithmic system of facial recognition. If here we have the evidence that targeted

²⁰⁸ UN Experts call for decisive measures to protect fundamental freedoms in China - <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=26006&LangID=E>

²⁰⁹ UN GA, Human Rights Council, *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression* (A/HRC/41/35/Add.2).

²¹⁰ WIEWIEÓROWSKI W., *A solution in Search of a problem?*, Available at: https://edps.europa.eu/press-publications/press-news/blog/facial-recognition-solution-search-problem_en, Accessed on 15th July 2020.

²¹¹ UN GA, Human Rights Council, *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression* (A/HRC/41/35/Add.2).

surveillance - applied to a certain type of people - creates problems such as self-exclusion, it suggests spontaneously to think that, even wanting to create a list of suspicious people to keep them under control, the effects may be the same. Therefore, if, on the one hand, a targeted list of individuals to watch is not created, the risk of discrimination and generalisation increases. Instead, on the other hand, if a possible list is created, either most of the people on the list will belong to vulnerable categories, or those same individuals will tend to exclude themselves from society (think about redeemed ex-convicts who are frightened by a surveillance system, even though they are not convicts anymore). Thus, in absence of a clear legal framework, those systems should be banned even acknowledging that good outcomes are possible and proven (e.g. surveillance systems to find missing children). To this idea of banning the system until new legislation, it could be argued that it is now too late. So, it might be useful for States to enable individual claims against both state and non-state actors²¹², so that those individuals, either being or not being on the watchlist can arise individual complaints.

Plus, the design of these technologies is very relevant. They must be designed to comply with human rights standard – which trace us back to question about universality of legislation, considering that this dilemma covers also human rights. Since the question about the universality of human rights remains open, we can state that when designing certain technologies, there are some principles that have to be respected depending on the historical time and place (today, for example, the principle of non-discrimination is dangerously at threat all over the world, therefore it should be of primary importance to protect it). Therefore, monitoring the design of those machines and systems is essential. Certainly, there is the need for those algorithms and system to be transparent as much as they can, in order to guarantee fairness. To verify the transparency, it would be useful to let companies be scrutinised from conception to implementation and any adverse impact on human rights must be sanctioned and changed²¹³.

Overall, there are many things -on the short term - that could be done in order to stop the perpetuation of bias. First of all, the banning of those technologies is quite relevant until a set of binding policies is developed everywhere. Issuing reports is of course fundamental, but lastly it is necessary to develop policies. When we talk about policy making, we also mean the use of existing policies that can be applied to new problems: the ICERD aforementioned is, for example, one of those. I thought as

²¹² UN GA, Human Rights Council, *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression* (A/HRC/41/35/Add.2), p. 16.

²¹³ MCNAMEE J., Report, Council of Europe, *High Level Conference – Governing Game changer – Impacts of Artificial intelligence development on human rights, democracy and the rule of law*, 26-27 February 2019, p. 5.

important to mention this convention because it represents the highest degree of defence of certain principles. However, adherence to respect certain values take place on a national basis and this does not guarantee respect for the rights by many countries which do not find themselves aligned with the values sustained by the UN. Nonetheless, it cannot be enough to apply a general principle such as non-discrimination to specific problems as algorithms. Secondly, interdisciplinarity and interculturality is essential. If a certain point of view about the world is imprinted into those mechanisms, the risk of discrimination grows exponentially. Diversity must be imprinted in those technologies and for this to happen, the workforce must be diverse. In fact, I think it should be made clear that e.g. monitoring must take place on different levels, and the mixture of different disciplines or the inclusion of women – who are excluded – must be welcomed. If an engineer is necessary in order to develop a machine or an algorithm, an ethics committee must be present alongside, able to recognize whether or not the algorithms, or the machine, will lead to further discrimination. Lastly, the fields in which those technologies are used could be restricted and reduced to the essentials. China's indiscriminate use of these technologies, for example, risks leading to the exasperation of an already authoritarian regime. If a model such as this were to start acting as an example, it would not only bring about a technological revolution, but also a change in the political, and ideological paradigm of thought. I strongly believe that it is precisely here, with regard to these issues, that we must think about “solutions” (if we want to call them so) in the long term.

3.4.2 Long term solutions

As Luciano Floridi said in an interview in this year 2020²¹⁴, we are experiencing a transformation that requires a rethinking of our concepts. Ideas such as identity, politics, society and, last but not least, human rights are being questioned. Therefore, we are talking about a re-ontology of modernity, a change in the conception of the being that, to date, is experiencing the hybridization with the machine and sees itself faced with a defrosting of its certainties²¹⁵. It seems that on the one hand, hybridization leads us to harmonization in a broad sense: harmonization as censorship, a Chinese term; or harmonization as unifying, making similar the different entities and identities present in the world. It is not clear in which direction this upheaval is bringing humanity. However, it is sure that is changing our lifestyle, conception of life and of other more “concrete” things such as rights. For example, the concept of privacy is expanding its boundaries to unimaginable levels, so much so that one wonders how much sense still makes to talk about privacy, as noted in Chapter 2. The struggles that have taken

²¹⁴ DANNA R., interview to Luciano Floridi “*Filosofia dell’Infosfera*”, Pandora Rivista, N° 1/2020.

²¹⁵ Ibid.

place in the name of equality - gender, race, ethnicity - are changing shape; they still have the taste of the struggle to defend mankind from mankind, but the background that accompanies them has changed, becoming more digital and apparently, and I stress apparently, less human. Even though this change is happening, it seems to be out of our reach, and we cannot regulate it, manage it, hold it in our hands. In fact, we perpetuate the same mechanisms of slavery through a machine. Remo Bodei says, “AI represents the overcoming of that division between arms and mind that founded the relationship of domination between master and servant, the risk is that old categories and therefore old hierarchies of power surreptitiously creep in and regenerate themselves in a different shape”²¹⁶. And that is exactly what is happening, old categories are simply being put back in a different form. If going back and uninvent is impossible, it seems necessary to act on two different levels. On the one hand, we need a cultural change, so we need to try to eradicate old systems of thought and to overcome racism, gender discrimination and so on. It might sound utopian to hope in a society where diversity is a gift, mostly because centuries of fights have not brought the required results. However, losing hope and leaving these machines in the hands of a few powerful people risks creating even more irreparable damage than many already committed. Man’s intelligence must reside on the one hand, now more than ever, in grasping the teachings that history has given us and re-adapting them to new paradigms, even if they seem to be unknown. On the other hand, humanity should try to get to know the so-called “unknown” and let the change could come from these very machines. If an algorithm proves to be impartial, if a robot does not look purely feminine and Western and, if FRS does not only target African Americans in America, maybe this can serve as a lesson for all humanity. Of course, this kind of change always occurs from the individual first and this is why, as said at the beginning, we need a cultural and ideological change (which, however, could start through those machines). This must certainly be accompanied by a real and well-known public debate which, to date, seems non-existent. Starting with the education of everyone, it is necessary to promote and involve all existing generations. Teaching how technology works and how it can be exploited positively is what is needed. For example, most people do not know about the bias in Amazon algorithms, even though everyone uses Amazon. Their algorithmic system for recruiting new employees, which, having received statistically many more male applications over the years, eliminated those received from women at the outset²¹⁷.

²¹⁶ PIGNATTI G., interview to Remo Bodei, “*Dominio e sottomissione. Schiavi, animali, macchine Intelligenza Artificiale*”, Pandora Rivista, N°1/2020.

²¹⁷ DASTIN J., *Amazon scraps secret AI recruiting tool that showed bias against women*, October 10, 2018, Reuters, Accessible on 24th July 2020, <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.

Overall, three main levels of action have therefore been identified. It is important to underline how those actions are seen as three straight lines that continue to intertwine and interchange. If they were to act independently on these three levels, the result would not be satisfactory. While there is a need for specific and targeted regulation, the key to these machines needs to be changed. They are scary, probably dangerous, but the only way to make them friends is to know them at best – mostly because they are still invented by us, the human being.

Although the machine is of human invention and therefore it is the result of natural intelligence, I think we can speak of an interchange between man and machine. If there is an interchange between the two, there is the possibility to benefit from it. As was said above, if we transmit fair values, considered so and shared by the most, we can make everything usable to a large part of the population. It remains true that technology creates division – for example, in the fieldwork, there are those who have the capacity to manage the technological system, people whose jobs require a lot of empathy, and on the other hand, there is a large mass of people who are destined to remain unemployed or to be pushed into low-skilled jobs. Nonetheless, I believe for technology to be a very useful means of educational expansion, if used correctly. Otherwise, *people will come to love their oppression*²¹⁸.

²¹⁸ POSTMAN N., *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*, Penguin Books; Anniversary edition, 2005.

THE UNEXPECTED PANDEMIC:

New Technologies during a global emergency

Unfortunately, when dealing with issues such as the ones brought by new technologies, it is impossible to neglect the latest events that have affected the entire globe. The recent Covid-19 pandemic has brought radical changes to our everyday life, to our worldview and to the way we monitor it. Some changes had to be made regarding the surveillance of individuals and, like everything related to the digital revolution, what is happening at the level of surveillance mechanisms is still quite unknown to most people. The decisions that have been taken by governments are affecting our economy, politics and culture²¹⁹. Facing a global emergency means, of course, taking decisions and passing them in hours rather than years, and this entails that the consequences are not as balanced as they would normally be. As we came across in the second chapter, the crisis of democracy is real and imminent, and the pandemic only brought this to the foreground.

In these months, democracy – and more than democracy, the democratic principles that guide many societies – are being sabotaged by a system that threatens to undermine our identity, putting it at the service of government and politics, without our consent. Which makes us question even more if democracy is still the best form of government.

Having access to all the data of a citizen should be a measure of extreme emergency, relegated only and exclusively to that particular moment in history. But who can assure us that this will be the case? Google, for example, having taken the helm very effectively, uses the concept of “sharing is caring” as its slogan²²⁰. The idea hidden behind this is to take control over the market, using the data to learn more about mind-body-environment interaction and ultimately to develop new products and services²²¹. To try to take control of the pandemic, Google, thanks to its map app, was able to create the Mobility-reports in 131 countries²²². Like Google, Facebook too has not missed the opportunity to extend its sphere of influence. Through social mapping and its “data for good”, Facebook wants to transform the future from uncertain to probable. These giants of technology, of course, leverage the emotionality of the average citizen, who will be positively impressed to know that thanks to data

²¹⁹ HARARI Y.N., *The world after coronavirus*, <<https://www.ft.com/content/19d90308-6858-11ea-a3c9-1fe6fedcca75>> Accessed on 17th July 2020.

²²⁰ NOSTHOFF A-V., MASCHEWSKI F., *La grande occasione*, Rivista Internazionale, 26 giugno 2 luglio, n 1364, anno 27.

²²¹ Ibid.

²²² Ibid.

collection systems and new technologies, they can save the world. What we experienced during the emergency is now the new normality, where surveillance technologies are allowed to penetrate our daily lives.

Additionally, FRS, has also an impact on everyday life and on the way the body is perceived. In fact, it does not take into account bodies. For example, traumatic experiences can have enormous impact on humans, at the point that their face can change.²²³ This is especially true when it concerns women who have a very high risk of experiencing traumatic situations due to gender-based violence²²⁴. This easily connects to the value of bodies in the age of the digital revolution. “Neoliberalism has made bodies the object of the market, seeking a post-organic life in which perfection could take concrete form through a human body. The catastrophic acceleration of the Anthropocene in the last thirty years testifies the harmful effects of the technical “everything is possible” that not only ignores, but tramples on the profound singularities of organic processes”²²⁵.

Technology needs to be guided through choices that take into accounts the interests of people and the social environment we live in²²⁶. What should concern us the most is that the increase in the usage of surveillance technology has not characterized just China, that installed CCTV camera points at the apartment door of those under a 14-day quarantine to ensure they don’t leave²²⁷ (which I find to be a particular case) but also countries as Israel, Singapore, Hong Kong and so on²²⁸. Surveillance took the chance to spread at the same speed of Coronavirus.

“(…) So, let's come to digital technology and say it now: there are no control technologies that are also “ethical”. Ethics develops in the relationship and is rooted in experience, it is not something that

²²³ GATES K., *Our biometric future: Facial Recognition Technology and the Culture of Surveillance*, New York University Press, 2011.

²²⁴ O’NEIL T., FLEURY A., FORESTI M., *Women on the move - Migration, gender equality and the 2030 Agenda for Sustainable Development*, (Swiss Agency for Development and Cooperation SDC, 2016) Available at <<https://www.odi.org/sites/odi.org.uk/files/resource-documents/10731.pdf>> accessed on 25 June 2020

²²⁵ BENASAYAG M., CANY B., DEL REY A., COHEN T., PADOVANO R., NICOTRA M., *Piccolo Manifesto in tempo di pandemia*, Available at: < <https://ilrasoiodioccam-micromega.blogautore.espresso.repubblica.it/2020/04/01/piccolo-manifesto-in-tempi-di-pandemia/>

²²⁶ Privacy Network, *The new normal. Riflessioni ai tempi del Covid-19*, Available at: <https://privacy-network.it/new-normal-riflessioni-covid19/>, Accessed on 17th July 2020.

²²⁷ <https://www.cnbc.com/2020/03/27/coronavirus-surveillance-used-by-governments-to-fight-pandemic-privacy-concerns.html>

²²⁸ KHARPAL A., *Use of Surveillance to fight coronavirus raises concerns about governments power after pandemic ends*, Available at: <https://www.cnbc.com/2020/03/27/coronavirus-surveillance-used-by-governments-to-fight-pandemic-privacy-concerns.html>.

can be embodied in code (no, free software does not guarantee the neutrality of the technique). Invasive control is never ethical”²²⁹.

Throughout the pandemic, we “experienced” a new collective solidarity, which had been lost prior to pandemic. However, it would be a mistake to believe that the collective character of the threat magically erases the disparities between bodies. Social class, gender, economic dominance, military violence or patriarchal oppression are all realities that situate our bodies differently. Let us not allow ourselves to be enchanted by the romanticising of confinement, which aims, by trumpeting, to make us forget these differences²³⁰. It is up to us to make it ethical as much as we can, and the only way we seem to have is through information and education.

If our task was already to try to understand the kind of revolution that is taking place, today more than ever, during this historical moment that will change the world as we know it, we have a duty to be responsible citizens. This must not just be relegated to a fortunate elite of the population, but to everyone. The wealthy part of the world must help the less wealthy and so on, so as to restore the balance of the world, which seems to be touching “the bottom”. The coronavirus crisis should have taught us that ‘nobody saves themselves’ without the others, because taken one by one, we are nothing²³¹.

²²⁹Gruppo Ippolita, *No, non esistono tecnologie di controllo etiche*, Available at: <<https://ilmanifesto.it/covid-no-non-esistono-tecnologie-di-controllo-etiche/>> Accessed on 17th July 2020.

²³⁰ BENASAYAG M., CANY B., DEL REY A., COHEN T., PADOVANO R., NICOTRA M., *Piccolo Manifesto in tempo di pandemia*, Available at: < <https://ilrasoiodioccam-micromega.blogautore.espresso.repubblica.it/2020/04/01/piccolo-manifesto-in-tempi-di-pandemia/>

²³¹ CECCOBELLI, D. interview to Luigi Di Gregorio “*Demopatia e crisi della sfera pubblica*”, Pandora Rivista, N°1/2020.

CONCLUSION

“Maybe digital technology will never be advanced enough in this century to dominate the economy, but it probably will. Perhaps technology will make all the demands of life so inexpensive that living well will be within everyone's reach and no one will worry about money, work, inequality of wealth or planning for old age.

I doubt very much that such a clear scenario will unfold. Instead, if we continue like this, we will probably enter a period of hyper-unemployment and the resulting political and social chaos. The result of the chaos is unpredictable, and we should not rely on it to plan our future²³²”.

Jaron Lanier, in his “Who owns the future” tries to predict what the future is holding for us and, in this passage, he states that chaos is – most probably – the right answer. In fact, in this research, it seems clear how taking a net position on the digital revolution and how it can be addressed would be presumptuous and inconclusive.

Instead, what we have proposed in this paper is an approach that stands in the middle, without easy excitement or dramatic despair. This is because – and during the work has emerged fully – the problem does not seem to be the technology itself, but the way this is used by the human being. There is no point in questioning the technology at one or the opposite extreme. The focus must move from technology to how we will face the transformation of our world. Asking what will happen to the world in the coming decades is a question that, to date, could be considered of secondary importance; it is rather relevant to ask ourselves what we want to be and how we are going to fit into the revolutions that will not be long in coming.

The approach we have used does not have sought a clear and unambiguous solution but often proceeded through questions, also because, given the extreme complexity of the subject, there was not another way to proceed. The introduction of new technologies in our daily life – more than it already is now – opens up unimaginable scenarios, and it is not possible to analyse them in their depths if we let ourselves be taken by surges of euphoria or, worse, by apocalyptic conclusions. What is hoped for today is therefore a continuous osmosis between the human being and the machine: the encounter between *humanitas* and *techne*²³³, in the end, appears the only way to follow in the coming years, aware that every choice must have a clear and certain shared basis and values. However, the human being is already thinking, taking decisions and reasoning through a constant interaction with

²³² LANIER J., *Who owns the future?*, First Round, ‘Motivation’, New York, Simon and Schuster 2013.

²³³ VINCENTI E., *Il «problema» del giudice-robot*, in A. CARLEO (a cura di), *Il Mulino*, Bologna, 2019, p. 123.

technology. Think about how we have changed the ways in which the individual decides where to go, what book to read or which movie to watch: if before we could rely only on our intuition, now we have a whole new world that helps us every day in every action we take. Therefore, there is much work to be done in this field. These pages are just a brief touch on a vast amount of problems concerning AI and new technologies.

We started outlining the fragility of one kind of political system, confronted to another one which seems to be much more powerful and dangerous at the same time. The power of silence that China has acquired, spread a veil of Maya over our eyes and made us think that efficiency – one of the greatest qualities of China – is the aim we set for this part of the globe too.

Consequently, what is happening is that we are forgetting all the battles that the generations that preceded us had to combat and the ones we are still fighting. Therefore, on the one hand we continue to fight for a society free of blind and square thought paradigms; and on the other hand, the revolution underway closes these possibilities. Moreover, intercultural dialogue between different countries must exist and must take the shape of a commutation of ideas and stimuli, rather than a conflictual exchange. If the ultimate goal of each country is supreme and subjugating power, then a positive development is not seen as possible. If China closes its doors to the idea of sharing the highly positive characteristics of its culture, a healthy transmission of thought cannot take place.

Equally, the discourse of gender inequality and racial discrimination must find its revolution at the roots of thought. Policies developed by local, national and international communities seem to look more like short-term solutions to appease minds rather than overturning a system of thought. While it is true and proven that cultural change requires decades to come, humanity cannot hope for history to change its course by its own. The machine used within the neoliberal context, purely individualistic ideology, breaks even more the bonds between individuals - if misused. The political context in which AI, and particularly the FRS is embedded is of fundamental importance. If the concept of democracy is falling apart, is easy to think that all the values within this concept are falling as well. Those values, such as gender equality or non-discrimination, are given and protected also thanks to the democratic principle and concept. Therefore, I found relevant to this research to touch upon the political context as well as the current violation of human rights where the technological revolution is taking place. International laws, treaties and conventions are one of the small steps towards change. However, international and national politics have flaws, therefore it will be essential - as has been pointed out - an approach that is as much possible shared, both within individual states and in the European context. As we have seen, there are already many documents and institutions dealing with this issue, trying to regulate and direct technological development. In particular, state authorities will have to respect the principles of respect for fundamental rights and non-discrimination. In addition, the principles of

transparency and fairness of the artificial systems and quality and safety of the new technologies must take primary importance. Certainly, it could be argued that the technological process should be stopped, or rather, should not have existed in this way. However, to proceed by utopias or denials of what surrounds us has proved, during the writing process, to be useless. It is certainly useful to try to imagine a future with different appearances, it is right to try to slow down the process and it seems reasonable to tell yourself and the whole world that it was not supposed to be like this. However, around this way of dealing with the problem, it has been understood that there must also be an active action, that somehow accepts the situation in which we find ourselves and tries to change it starting from what it is and not from ‘what it could have been’.

The gradual introduction of artificial systems into our lives will be able to have a well-controlled path and – above all – shared by as many parties as possible, avoid an uneven application that would not have the strength to bring significant advantages. Hence, alongside a technological revolution, a cultural revolution must walk. The paradigms of thought now in force appear to be unsuitable for the historical moment we are living, thus proving unjust and unfair solutions. At the same time, laws and regulations must transmit these changes - technological and cultural - and regulate new problems with a mixture of old and new paradigms. As has been argued, the machine will itself serve *ex post* for review some human’s decision, pointing out flaws, presenting different and alternative solutions, etc.²³⁴. In this way the symbiosis between man and machine, between nature and artifice can be virtuous, devoid of the extolled myths of a machine that dethrones man in every way. The machine thus becomes an essential aid and – with the passing of the years – its importance will be felt more and more in a future that will see the proliferation of the technique but which – and it is said with certainty – will not put an end to the anthropocentric view of the world. Technology must not become the *Leviathan* of our century.

²³⁴ A. PUNZI, Judge in the machine. E se fossero le macchine a restituirci l’umanità nel giudicare? In A. CARLEO (a cura di) *Decisione robotica*, Il Mulino, Bologna, 2019, p. 330

BIBLIOGRAPHY

Articles

BENASAYAG M., CANY B., DEL REY A., COHEN T., PADOVANO R., NICOTRA M., *Piccolo Manifesto in tempo di pandemia*, Available at: <https://ilrasoiiodiocammicromega.blogautore.espresso.repubblica.it/2020/04/01/piccolo-manifesto-in-tempi-di-pandemia/>

BOTSMAN R., *Big Data meets Big brother as China moves to rate its citizens*, 21 October 2017, <https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion>, Accessed on 18th July 2020.

BUCKLEY C., MOZUR P., *How China uses High-Tech surveillance to Subdue Minorities*, Available at: <https://www.nytimes.com/2019/05/22/world/asia/china-surveillance-xinjiang.html>

BUOLAMWINI J., *We Must Fight Face Surveillance to Protect Black Lives*, Available at: <https://onezero.medium.com/we-must-fight-face-surveillance-to-protect-black-lives-5ffcd0b4c28a>, Consulted on 30th June 2020.

BUOLAMWINI J., *Facial recognition software is biased towards white men, researcher finds*, Available at: <https://www.media.mit.edu/articles/facial-recognition-software-is-biased-towards-white-men-researcher-finds/>, Consulted on 2nd July 2020.

BUZZANO G., *The age of surveillance capitalism by Shoshana Zuboff*, Pandora Rivista, N° 1/2020.

CECCOBELLI, D. interview to Luigi Di Gregorio “*Demopatia e crisi della sfera pubblica*”, Pandora Rivista, N° 1/2020

Clarks Legal, Case analysis, https://www.clarkslegal.com/Blog/Post/Case_Analysis_R_Bridges_v_CCSWP_and_SSHD_2019_EWHC_2341 Accessed on 12th July 2020.

Contributor CSM, *Facial Recognition Security Technology: The Facts*, 1st May 2020, <https://citysecuritymagazine.com/security-technology/facial-recognition-security-technology-the-facts/> Accessed on 19th July 2020.

DANNA R., interview to Luciano Floridi “*Filosofia dell’Infosfera*”, Pandora Rivista, N° 1/2020.

DASTIN J., *Amazon scraps secret AI recruiting tool that showed bias against women*, October 10, 2018, Reuters, Accessible on 24th July 2020, <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.

Data Sanitization, Available at: <https://www.datasanitization.org/data-sanitization-regulations/>

Department of Digital, Culture, Media & Sports, Available at: <https://www.gov.uk/government/news/government-to-strengthen-uk-data-protection-law>

DE VERE C., *Facial Recognition in public places*, Available at: <https://www.infolaw.co.uk/newsletter/2019/10/facial-recognition-public-spaces/>, Accessed on 13th July 2020.

GOLDMAN D., *Worst year for jobs since '45*, Available at: https://money.cnn.com/2009/01/09/news/economy/jobs_december/: Consulted 8th June 2020.

Gruppo Ippolita, *No, non esistono tecnologie di controllo etiche*, Available at: <https://ilmanifesto.it/covid-no-non-esistono-tecnologie-di-controllo-etiche/> Accessed on 17th July 2020.

HARARI Y.N., *The world after coronavirus*, <https://www.ft.com/content/19d90308-6858-11ea-a3c9-1fe6fedcca75> Accessed on 17th July 2020.

HILL K., *Wrongfully accused by an algorithm*, 2020, Available at: <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>, Accessed: 22 June 2020.

HOOGHIEMSTRA T., *Informational Self-Determination, Digital Health, d New Features of Data Protection*, Available at: https://edpl.lexxion.eu/data/article/14259/pdf/edpl_2019_02-007.pdf.

Information Commissioner's Opinion, *The Use of live Facial recognition technology by law enforcement in public places*, Available at <https://icosearch.ico.org.uk/s/search.html?query=R+%28Bridges%29+v+South+Wales+Police%2C&collection=ico-meta&profile=default>.

KARI P., *San Francisco is First US City to ban police use of Facial Recognition Tech*, Available at: <https://www.theguardian.com/us-news/2019/may/14/san-francisco-facial-recognition-police-ban>

KHARPAL A., *Use of Surveillance to fight coronavirus raises concerns about governments power after pandemic ends*, Available at: <https://www.cnn.com/2020/03/27/coronavirus-surveillance-used-by-governments-to-fight-pandemic-privacy-concerns.html>.

JACQUES M., *Can the West's democracy survive China's rise to dominance?*, Available at: <https://www.economist.com/open-future/2018/06/14/can-the-wests-democracy-survive-chinas-rise-to-dominance>, Consulted 20th June 2020.

O'NEIL T., FLEURY A., FORESTI M., *Women on the move - Migration, gender equality and the 2030 Agenda for Sustainable Development*, (Swiss Agency for Development and Cooperation SDC,

2016) Available at <<https://www.odi.org/sites/odi.org.uk/files/resource-documents/10731.pdf>> accessed on 25 June 2020.

OVIDE S., *The Real Dangers of Surveillance*, Available at: <https://www.nytimes.com/2020/06/12/technology/surveillance-protests-hong-kong.html>.

Privacy Network, *The new normal. Riflessioni ai tempi del Covid-19*”, Available at: <https://privacy-network.it/new-normal-riflessioni-covid19/>, Accessed on 17th July 2020.

SIMONITE T., *The Best Algorithms Struggle to Recognize Black Faces Equally*, Available at: <https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally/>, Consulted on 28th June 2020.

SINGER N., METZ C., *Many Facial Recognition Systems Are Biased, Says U.S. Study*, <https://www.nytimes.com/2019/12/19/technology/facial-recognition-bias.html>, Accessed on 27th June 2020.

STUART F., *How Facial Recognition technology is bringing surveillance capitalism to our streets*, 2019, Available at: <https://www.opendemocracy.net/en/oureconomy/how-facial-recognition-surveillance-capitalism-streets/>

VALLOR S., *Lessons From Isaac Asimov’s Multivac*, 2017, Available at: <https://www.theatlantic.com/technology/archive/2017/05/lessons-from-the-multivac/523773/> Consulted 25th April 2020.

VINCENT J., *Twitter taught Microsoft’s AI chatbot to be a racist asshole in less than a day*, 2016, March 24, The Verge, <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>, Accessed on 19th July 2020.

VINCENT J., *Gender and Racial bias found in Amazon’s facial recognition technology (again)*, Available at: <https://www.theverge.com/2019/1/25/18197137/amazon-rekognition-facial-recognition-bias-race-gender>, Consulted on 14th June 2020.

WEST S.M., WHITTAKER M., CRAWFORD K., *Discriminating Systems: Gender, Race and Power in AI*, AI Now Institute, 2019, Available at: <https://ainowinstitute.org/discriminatingystems.html>, Accessed on 19th July 2020.

Books and Journals

BARTLETT J., *The people vs tech. How the internet is killing democracy*, Ebury Press, 2018.

BATACHARIA B., LEVY D., CATIZONE R., KROTOV A., WILKS Y., (1999) *CONVERSE: A Conversational Companion*, In: Wilks Y. (eds) *Machine Conversations. The Springer International Series in Engineering and Computer Science*, vol 511. Springer, Boston, MA.

BAUMAN Z., *Liquid Modernity*, Polity Press, 2000.

BENASAYAG M., *La tirannia dell'algoritmo*, Vita e Pensiero, 2020.

BONARINI A., *Sistemi Fuzzy*, in *Mondo digitale*, 1, 2003.

BOUCHER P., *How artificial intelligence works*, Briefing of the European Parliament, March 14th, 2019

BUOLAMWINI J., GEBRU T., *Gender Shades, Intersectional Accuracy Disparities in Commercial Gender Classification*, Sorelle A. Friedler and Christo Wilson, 2018.

BVerfGE 65, 1 - Volkszählung Judgment of the First Senate of 15 December 1983 on the mündliche Hearing of 18 and 19 October 1983 - 1 BvR 209, 269, 362, 420, 440, 484/83 in the proceedings über the constitutional complaints.

CARLUCCI AIELLO L., DAPOR M., *Intelligenza artificiale: i primi 50 anni*, in *Mondo digitale*, 2004.

COSTANZI C., *La matematica del processo: oltre le Colonne d'Ercole della giustizia penale*, in *Questione giustizia*, trimestrale n. 4, 2018.

DE VOS M., *The European Court of Justice and the march towards substantive equality in European Union anti-discrimination law*, *International Journal of Discrimination and the Law* 2020, Vol. 20(1).

EINSTEIN A., *Sidelights on relativity*, 1922.

Freie Universität Berlin - Team of scholars at surveyed 2,200 citizens, “*Study: More than two thirds of Chinese take a positive view of social credit systems in their country*”, N° 198/2018 from July 23, 2018.

GATES K., *Our biometric future: Facial Recognition Technology and the Culture of Surveillance*, New York University Press, 2011.

HARARI Y.N., *21 Lessons for the 21st Century* (2018), trad.it. di M. PIANI, Bompiani editore, 2018.

HAUGELAND J., *Artificial intelligence: the very idea*, MIT press, Cambridge, 1985.

International Journal of Scientific and Research Publications, *A survey for Facial Recognition technology*, Volume 6, Issue 7, July 2016.

LANIER J., *Who owns the future?*, First Round, 'Motivation', New York, Simon and Schuster 2013.

LIN P., ABNEY K., BEKEY G., *Robot ethics: Mapping the issues for a mechanized world*, in *Artificial intelligence*, 175, 2011.

LIU HW., LIN CF., CHEN YJ., *Beyond State v Loomis: artificial intelligence, government algorithmisation and accountability*, *International Journal of Law and Information Technology*, Volume 27, Issue 2, Summer 2019.

LORBER J., FARRELL S. A., *The social construction of gender*, SAGE Publications, 1991.

MAC SÍTHIGH D., SIEMS M., *The Chinese Social Credit System: A model for Other Countries?*, *Modern law review*, 2019.

MCCARTHY J., MINSKY M., ROCHESTER N., SHANNON C., *A proposal for the Dartmouth summer research project on artificial intelligence*, 1955.

MESSETTI G., *Nella testa del Dragone*, Mondadori, 3 Marzo 2020.

NOBLE S. U., *Algorithm of oppression. How search engines reinforce racism*, New York University Press, 2018.

NOSTHOFF A-V., MASCHEWSKI F., *La grande occasione*, *Rivista Internazionale*, 26 giugno 2020, n 1364, anno 27.

OSWALD M., GRACE J., URWIN S., BARNES G. C., *Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality*, *Information & Communications Technology Law*, 2018, 27:2.

PIGNATTI G., interview to Remo Bodei, "*Dominio e sottomissione. Schiavi, animali, macchine* *Intelligenza Artificiale*", Pandora Rivista, N°1/2020.

POSTMAN N., *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*, Penguin Books; Anniversary edition, 2005.

PUNZI A., Judge in the machine. E se fossero le macchine a restituirci l'umanità nel giudicare? In A. CARLEO (a cura di) *Decisione robotica*, Il Mulino, Bologna, 2019.

QUING BA C., *Chinese Constitutional law*, Available at: https://heinonline.org/HOL/Page?collection=journals&handle=hein.journals/brac1j26&id=77&men_tab=srchresults Consulted 15th July 2020.

ROUVROY A., POULLET Y., *The right to informational self-determination and the value of self-development. Reassessing the importance of privacy for democracy*, in *Reinventing Data Protection?*, ed. Serge Gutwirth et al. (Dordrecht: Springer, 2009).

SARTOR G., *Intelligenza artificiale e diritto. Un'introduzione*, Giuffrè, Milano, 1996.

SHAW J., *Making Evil: The Science Behind Humanity's Dark Side*, Canongate Books Ltd, 2019.

SOMALVICO M., AMIGONI F., SCHIAFFONATI V., *Intelligenza artificiale*, 2019.

STRITTMATTER K., *We have been harmonized. Life in China's surveillance State*, Costum House, 2020.

TURING A.M., *Computing machinery and intelligence*, in *Mind*, 59, 1950.

VALERA, L., *Post-humanism: beyond humanism?*, Institute of Scientific and Technological Practice, Università Campus Bio-Medico di Roma, Copyright Cuadernos de Bioética, 2014.

VINCENTI E., *Il «problema» del giudice-robot*, in A. CARLEO (a cura di), *Il Mulino*, Bologna, 2019.

ZUBOFF S., *The age of surveillance capitalism. The fight for a human future at the new frontier of power*, Public Affairs New York, March 2020.

WARWICK K. *Artificial intelligence - The basics*, Routledge, 2012.

WARWICK K., SHAH H., *Turing's imitation game – Conversation with the unknown*, Cambridge University Press, 2016.

WOODWARD J. D., HORN JR. C., GATUNE J., THOMAS A., *Biometrics: A Look at Facial Recognition*, Rand Publishing, 2003.

Council Decision (EU) 2013/32 of 9 April 2019 authorising Member States to ratify, in the interest of the European Union, the Protocol amending the Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data [2013] OJ L180/60.

Constitution of People Republic of China, adopted 4th December 1982.

Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, 4 November 1950) (ECHR).

Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast) OJ L 204, 26.7.2006.

EU Charter of Fundamental Rights: Charter of Fundamental Rights of the European Union [2000] OJ C364/1.

European Commission, High level expert group on Artificial Intelligence, set up by the European commission, Ethics guidelines for trustworthy AI, 8th April 2019.

GOODMAN B., *Discrimination, Data Sanisation and Auditing in the European Union's General Data Protection Regulation*, Heinoline, 2016.

Handbook on European non-discrimination law 2018 edition (FRA European Union Agency for Fundamental Rights, 2018) 18 available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-handbook-non-discrimination-law-2018_en.pdf, Accessed on 5 June 2019.

MCNAMEE J., Conference Report, HIGH-LEVEL CONFERENCE, *Governing the Game Changer – Impacts of artificial intelligence development on human rights, democracy and the rule of law*, conference co-organised by the Finnish Presidency of the Council of Europe Committee of Ministers and the Council of Europe, 26 – 27 February 2019, Helsinki, Finland – Finlandia Hall.

MCNAMEE J., Report, Council of Europe, *High Level Conference – Governing Game changer – Impacts of Artificial intelligence development on human rights, democracy and the rule of law*, 26-27 February 2019.

Protocol No. 12 to the convention for the protection of human rights and fundamental freedoms (adopted 4 November 2000).

R (Bridges) v CCSWP and SSHD <<https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf>> Accessed on 10th July 2020.

State Council Notice concerning Issuance of the Planning Outline for the Construction of a Social Credit System (2014-2020), *Planning Outline for the Construction of a Social Credit System*, (2014-2020),

Available at <https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/> Accessed on 20th June 2020.

State v. Loomis, Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing, in Recent cases, Harv. L. Rev., 130, 2017.

State v. Loomis, 881 N.W.2d 749 (2016) the Courts states: «However, the COMPAS risk assessment does not predict the specific likelihood that an individual offender will reoffend. Instead, it provides a prediction based on a comparison of information about the individual to a similar data group».

STUPP C., *EU Plans Rules for Facial-Recognition Technology*, Available at: <https://www.wsj.com/articles/eu-plans-rules-for-facial-recognition-technology-11582219726>, Accessed on 10th July 2020.

UN Experts call for decisive measures to protect fundamental freedoms in China - <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=26006&LangID=E>

UN GA, Human Rights Council, *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression* (A/HRC/41/35/Add.2).

UN, GA, Res 217 A(III), *Universal Declaration of Human Rights* (adopted 10 December 1948) (UDHR).

UN, GA, Res 2200 A(XXI), *International Covenant on Civil and Political Rights* (adopted 16 December 1966) 999 UNTS 171 (ICCPR).

UN GA, Res 2200 A(XXI), *International Covenant on Economic, Social and Cultural Rights*, (adopted 16 December 1966), United Nations, Treaty Series, vol. 993, 3 art 3.

UN, GA, International Convention on the Elimination of All Forms of Racial Discrimination, Adopted and opened for signature and ratification by General Assembly resolution 2106 (XX) of 21 December 1965, art 1.

United Nations, Human Rights Office of the High Commissioner, The United Nations Human Rights Treaty System, Fact Sheet No.30, Rev 1, New York and Geneva, 2012.

United Nations Office of the High Commissioner of Human Rights, 'Women's Human Rights and Gender Equality', 2019. Available at <https://www.ohchr.org/EN/Issues/Women/WRGS/Pages/WRGSIndex.aspx>, Accessed on 29 June 2019.

United Nations, General Assembly, *Report on the Promotion and Protection of the right to freedom of opinion and expression*, A/73/348 (29 August 2019).

United Nations, Human Rights Office of the High Commissioner, The United Nations Human Rights Treaty System, Fact Sheet No.30, Rev 1, New York and Geneva, 2012.

United Nations, <<https://www.un.org/ruleoflaw/thematic-areas/human-rights/equality-and-non-discrimination/>> Accessed on 15th July 2020.

United Nations Secretary-General's strategy on New Technologies, <https://www.un.org/en/newtechnologies/> Accessed on 10th July 2020.

Other Sources

BORGESIOUS F. Z., *Discrimination, artificial intelligence and algorithmic decision-making*, Council of Europe, 2018.

- BOSCO F., CREEMERS N., FERRARIS V., GUAGNIN D., KOOPS B.-J., *Reforming European data protection law*, Law, government and technology series, Springer 2011.
- COLLETT C., DILLON S., *AI and Gender. Four Proposals for Future Research*, University of Cambridge, 2019.
- CROZIER M., HUNTINGTON S.P., WATANUKI J., *The crisis of democracy. Report on the Governability of Democracies to the Trilateral Commission*, New York University Press, 1975.
- Economist Intelligence Unit - A report by the Economist Intelligence Unit - Democracy index 2019.
- FELDSTEIN S., *The Global Expansion of AI Surveillance*, Carnegie Endowment for International Peace, 2019.
- FISCHER-HÜBNER S., HOOFNAGLE C., Krontiris I., RANNENBERG K., WAIDNER M., Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2468200.
- GATES K., *The past perfect promise of Facial Recognition Technology*, Institute of Communication Research, University of Illinois at Urbana-Champaign, 2004.
- LANGDON W., *Do artifacts have politics?*, MIT Press on behalf of American Academy of Arts and Science, 2009.
- MCCARTHY J., *Programs with common sense*, Stanford University, Stanford, 1959.
- PATO J.N., MILLETT L. I., *Biometrics Recognition: Challenges and Opportunities*, National Research Council of the National Academies, The National academy Press, Washington D.C., 2010.
- THOMI L. B., *The Impact of Biometric Systems at EU Outside Borders on the Human Rights of Irregular Migrating Women*, Global Campus of Human Rights, 2018/2019.
- TZANO M., Data protection as a fundamental right next to privacy? ‘Reconstructing’ a not so new right.
- ZANFIR-FORTUNA G., *The general data protection for regulation, analysis and guidance for US Higher Education institutions*, Ashleigh Imus editor, 2020.
- WIEWIEÓROWSKI W., *A solution in Search of a problem?*, Available at: https://edps.europa.eu/press-publications/press-news/blog/facial-recognition-solution-search-problem_en, Accessed on 15th July 2020.

