Università della Calabria

Department of Computer Engineering, Modelling, Electronics and Systems

Ph.D. in Information and Communication Technologies (XXXVI)

Thesis

Ethical AI: definition of the techno-legal rules to oversee the decisions of the automaton.

Advisor
Prof. Irina Trubitsyna

Irina Trubitsyna
25.02.2024 21:02:10
GMT+01:00

Co-Advisor
Prof. Pasquale Laghi

PASQUALE LAGHI
25.02.2024
22:20:41
GMT+01:00

Ph.D. Coordinator
Prof. Giancarlo Fortino

GIANCARLO
FORTINO
26.02.2024
06:41:01
GMT+00:00

Candidate
Maurizio Ferrari

Reg. n. 226125

MAURIZIO FERRARI
25.02.2024 22:26:31
GMT+01:00

Academic Year 2023/2024

INDEX

INTRODUCTION

The thesis work aims to analyze - within the current topics of great interest in artificial intelligence - the problem of defining norms and rules that can be used by the robot/automaton in making decisions. A problem that has roots dating back, so much so that the so-called three laws of robotics were developed by Asimov in 1950[1], very general principles that express basic ethical options, but appear completely insufficient to regulate the specificity of the different domains.

The need to give substance to specific systems of rules arises from the pace of technological progress, which continually gives rise to new situations worthy of attention[2], which cannot always be framed within an already existing system of rules. There is a particularly strong need, then, to give space to an anthropocentric vision which is not limited to allowing the a posteriori discernment of the operations of computer systems in general, and of automatons in particular, predicting that regulatory compliance also plays a role priors. This would allow, already in the conception, design and implementation phases of the IT systems, the identification of specifications that take into account ex ante the ethical and regulatory constraints applicable to the activities that they will/can carry out, ensuring that the engineering of the software is aimed not only at the needs of application effectiveness and efficiency, but also at guaranteeing its compliance with the aforementioned ethical and regulatory constraints.

Legal IT, from the point of view of jurists, up to now has mainly dealt with the topic of the possible uses and applications of IT in the world of law, however complex, uniquely oriented within the so-called human sciences. Technological innovation, however, imposes a new approach, presenting itself not only as a complicating factor of purely

[1]I. Asimov, I, Robot, New York, 1950. The so-called three laws of robotics are the following: First Law: A robot cannot harm a human being nor can it allow a human being to receive harm. Second Law: A robot must obey orders given by humans as long as such orders do not contravene the first law. Third law: A robot must protect its existence as long as this does not conflict with the first and second laws. Then there is the so-called Zero Law - aimed at resolving potential conflicts between the other three - according to which: A robot cannot cause harm to humanity, nor can it allow humanity to suffer harm due to its failure to intervene .

[2]Think of the example of the self-driving car, now a reality, which in dealing with the risks of road traffic "must make decisions", not only in compliance with the rules of the road, which would prove insufficient to avoid overwhelming users of the road who do not act in compliance with those rules, but considering ethical and legal aspects: how can the car decide not to hit a pedestrian who is crossing the road outside the pedestrian crossing?

legal problems, but as a new place in which the law is called to reside: cyberspace[3]Upon closer inspection, as has recently been noted in doctrine[4], it constitutes the bed in which the legal rules are destined to conform to the architecture of this new "artificial" environmental context in which the dimension of "being" coincides with that of "functioning".

Furthermore, the need to create organizational and regulatory compliance forms that allow operators in all sectors involved in the use of cyberspace to constantly verify the regulatory compliance of the concrete actions implemented in order to avoid (or in any case limit) reputational damage and risks associated with compliance with legislative, regulatory or self-regulation rules.

Inevitably, new risk situations are taking shape, which are the object of the protection purposes that support the legal system of civil liability, which can be placed in the conceptually unitary, but in concrete reality articulated and complex, category of "technological risk". . This is a category of risks closely connected to the pace of technological progress, sharing a common technical matrix with IT systems which, in its neutrality, can be used to strengthen security levels, just as it can be used to circumvent them. . The risk belonging to this category can be minimized, managed, but not eliminated[5]. The category itself turns out to be eclectic and changeable, so it is not possible to gain an exhaustive understanding of it, for the purposes of legal qualification, except through a cognitive investigation which, not limiting itself to mere legal aspects, has its roots in the evolution technical aspects and technological development.

In traditional legal categories, the civil liability system is described as a regulatory system that carries out various functions that are modeled on the varied protection needs that the factual reality brings out in correspondence with the risks that, either due to natural occurrences (such as the occurrence of disasters more or less predictable), or due to artificially determined events (when the risk originates, directly or indirectly, from

[3]It should be noted that, among the first scholars who authoritatively dealt with the topic, N. Irti, *Le categorie giuridiche della globalizzazione*, in *Riv. dir. civ*., 2002, p. 626; ID., *Norma e luoghi. Problemi di geo-diritto*, Roma-Bari, 2006, p. 16, ne ha dato la definizione di “non luogo”, to mark the characteristics of an area separated from space and time.

[4]Cfr. P. Laghi, *Cyberspazio e sussidiarietà*, Napoli, ESI, 2015.

[5]According to S. Garfinkel, G. Spafford, Web Security & Commerce, Sebastopol, O'Reilly, 1997, p. 9, «the only truly secure computer system is one that has been turned off, placed inside a concrete block and sealed in a room surrounded by security guards».

human actions), constitute a risk of "unjust damage"[6]or damage (pecuniary or otherwise) to subjective situations worthy of protection.

One of the most debated issues among jurists concerns the configurability of hypotheses of "strict liability", normally admitted in the tradition of Common Law, but accepted with difficulty in that of Civil Law[7]. In such cases the protection needs, in application of an efficient Economic Analysis of Law criterion[8], ignore the identification of fault as the basis of responsibility, to rise to the configuration of "liability for occurrence" (based on the principle ubi commoda, ibi incommoda) as opposed to "liability for conduct". A sector particularly affected by the logical-legal problems of "no-fault" liability is that of road traffic damage, for which the Italian civil code sets up "presumed fault" liability mechanisms[9], whose regulatory provision is inspired by the aim of making the system of protection of injured parties efficient, removing it from the complications deriving from the general criterion of attributing the damage (and its compensation) to the person whose fault can be demonstrated[10] .

In light of technological development and the growing types and types of risky situations that can be included in the "technological risk" category, the civil liability system needs to be revisited, with particular reference to the fact of liability that occurs in the context of activities carried out with aid of artificial intelligence, just as legal concepts that are no longer adequate, such as that of legal subjectivity, must be revisited with respect to the widespread use of algorithms in the establishment, modification and extinction of legal relationships.

The breadth and complexity of such a work, affecting the most varied sectors of the legal system that suffer the impact of technological evolution, necessarily transcends the economy of the present research project, whose aim is to analyze the problem of definition of norms and rules that can be used by the robot/automaton in making

[6]The expression is found verbatim in the art. 2043 of the Italian civil code, a rule that updates the principle, dating back to the Lex Aquilia de damno (year 286 BC), traditionally expressed in the general precept of neminem laedere.

[7]In Italy, among the first to point out the possibility of an underlying reading, in North American style, of some provisions of the civil code which are only apparently coherent with the system of civil liability anchored to the detection of the agent's fault, was P. Trimarchi, Risk and objective liability, Milan, Giuffrè, 1961.

[8]See R. Posner, Economic Analysis of Law, edit. WolterKluwer Law & Business (January 3, 2014).

[9]The art. 2054 of the Civil Code establishes the presumption of fault of the driver of the vehicle in the event of a collision with a pedestrian, and the equal contribution of fault of all drivers in the event of a collision between vehicles.

[10]For a view from a Common Law perspective, see G. Calabresi, The Costs of Accidents, Yale University Press, 1970.

decisions, through the identification of forms and types of techno-legal rules, which can allow the engineering of software "under regulatory constraints", relevant and usable already in the phases of conception, design and implementation of information systems.

The main aim of the research project is to analytically discuss the topic of the normative self-conformation of autonomous agents, in order to verify which tools and methods can allow the subsidiary interaction between legislative factors and project settings in the development of the techno-legal rule.

We will therefore proceed with framing the state of perception of the situations relating to the category of "technological risk", first of all keeping in mind the "European Union Guidelines for anthropocentric and reliable artificial intelligence"[11], who specifically, as regards the topic of the research project, believe it is essential that the algorithms underlying trustworthy AI are safe, reliable and sufficiently robust to deal with errors or inconsistencies during all phases of the systems life cycle computer scientists.

The centrality of human dignity and freedom, even and above all when mechanized systems and algorithms are involved, implies the necessary prevalence of man over artificial autonomy and this means that, right from the design phase, the system must be set up in compliance with the requirements of reliability which the robot/automaton must be able to match in every single choice even if unpredictable and even if "tragic"[12].

The techno-juridical rule must be conceived through logical-operational procedures of convergence between purely technical rules, ethical principles and legal norms - the application of which would be essential in the event that the decision of the robot/automaton were revealed a posteriori to be a harbinger of damage - precisely at the aim of avoiding that the "solution" of the case must be solely and necessarily legal, as such entrusted to a posthumous evaluation in terms of civil liability with the result of being able to guarantee only compensatory remedies.

The expected results are the identification of methods and criteria which, with reference to the specificity of concrete situations, allow us to arrive at a specific techno-legal rule which determines its implementation in IT systems already in the design phase, in order

[11]They provide that, to qualify as anthropocentric and reliable, AI must satisfy three fundamental components: 1) be in compliance with the law, and in particular with the EU Treaties, the Charter of Fundamental Rights of the European Union and the international human rights; 2) respect ethical principles; 3) be solid (see “Ethical guidelines for trustworthy AI” Communication from the Commission of 8 April 2019, COM-2019/168).

[12]On the topic of tragic choices, framed within the economic analysis of law, cf. P. Bobbit-G. Calabresi, The tragic choices, New York, 1978.

to use its potential also in order to provide responses to unpredictable risk situations in which the machine - with a view to preventing the occurrence of the harmful event rather than having to then compensate it when the damage has occurred - will necessarily have to make the "choice right."

CHAPTER I

Artificial Intelligence

The term “artificial intelligence” was first used by John McCarthy, assistant professor of mathematics at Dartmouth College in Hanover[13], in 1955. The scholar had organized a summer seminar called "Summer Research Project on Artificial Intelligence", bringing together experts from all over the world with a single objective: to develop a machine capable of simulating human intelligence.[14]

These were the years immediately following the Second World War and everyone still had in mind the machine "The Bombe", the forerunner of modern computers, which had been able to beat "Enigma", the device used by the Germans to encrypt their communications.

It is no coincidence, in fact, that, in that same year, Alan Turing, inventor of "The Bombe" and founding father of the entire subject, published the article "Computing machinery and intelligence"[15] asking the fundamental question: “Can machines think?”

Are machines capable of thinking? Or rather, is it possible to codify human thought in a binary system, typical of programming languages? These were the questions that inspired the participants at the Dartmouth conference and subsequent generations of researchers. As already anticipated in the introduction, the particularly controversial answers will be analyzed later in the discussion.

Before continuing with the description of the major inventions over the years, it seems appropriate to briefly analyze the main approaches that have directed the work of scholars in the field of artificial intelligence.

In particular, the main schools of thought are the cognitive model and the connectionist model, influenced by the progress of the various philosophical-anatomical disciplines that have been concerned with deepening the study of the human mind, such as psychology and neurobiology.

The cognitive model approach, also called the "symbolic" approach, has long been

[13]Kaplan J., Artificial intelligence: a guide to the near future, LUISS University Press, 2017, 37.
[14]Haenlein – Kaplan, A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence, California Management Review 61, no. 4 (2019), 7.
[15]Turing, I.—Computing Machinery and Intelligence, Mind LIX, no. 236 (January 1950): 433- 460

considered the "classic" model of artificial intelligence.[16]

The main characteristic of this approach was the conception of human thought as "conducted on mental representations, i.e. the symbols of the so-called language of thought, capable of sharing some syntactic and semantic properties with the expressions of natural languages."[17]

In particular, therefore, cognitivists aimed to reproduce human intelligence from a serial computer through the development of programs that simulated human reasoning in a symbolic language.

As you will see, the main applications of this approach have led to the creation of "inductive" machines, which exploit initial knowledge to arrive at a specific conclusion.

Differently, supporters of the concession model started from the fundamental assumption according to which a machine would be able to "take advantage" of human intelligence only after understanding its fundamental mechanisms.

Their efforts, in fact, have focused on the development of the so-called "neural networks", inspired by the structure of the human brain and the numerous "interconnections" between its basic units, the neurons.

In this regard, the objective was to replace the sequential computational architectures of cognitivists, exploiting an architecture more congenial to human brain activity, carried out in parallel and not in sequence.[18]

In this sense, the work carried out in 1943 by Warren McCulloch and Walter Pitts, respectively an American neurophysiologist and an American mathematician, was decisive, culminating in the publication of the conclusions reached following a study on the composition of the human brain and, specifically, on the neurons.[19]

The two scholars hypothesized, in fact, the possibility of programming a machine capable of functioning on the basis of a structure similar to that of the human nervous system which, as we will see, inspired the creation of the first neural networks.

---

[16]Elisa Gambetti, Riccardo Buscaroli, Federico Chesani, Fiorella Giusberti, Daniela Loreti, Paola Mello, Artificial intelligence and cognitive psychology compared in the solution of mathematical games, in Intelligent systems, Quarterly journal of cognitive sciences and artificial intelligence, 2/2020.

[17]Luciano Arcuri, Social neuroscience: a possible path to overcome the difficulties of cognitivism, in "Social Psychology, Rivista quadrimestrale", 1/2006.

[18]Paul Smolensky, Connectionism between symbols and neurons, translated by Frixione M., Marietti/Cambridge University Press, 1992.

[19]Warren S. McCulloch and Walter Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity, in *The Bulletin of Mathematical Biophysics*5, no. 4, 1943.

Once the analysis of the main features of both models has been completed, it is therefore possible to begin the reconstruction of the development of the AI.

The years immediately following the Dartmouth conference were characterized by a great commitment to programming which contributed to the realization of the first steps in the development of intelligent machines, despite a not particularly high level of technical knowledge and a general distrust on the part of investors for the financing of research.

As already explained previously, the cognitivist model played a leading role in the early days of artificial intelligence.

In fact, after an initial impetus which materialized with the "Perceptron"[20]by Rosenblatt, considered one of the first attempts to put McCulloch and Pitts' theories into practice and create a neural network capable of analyzing and recognizing images,[21]the connectionist model was quickly set aside and revived only in the 1980s.

Decisive, in this sense, was a harsh criticism by Marvin Minsky, one of the greatest authorities on intelligent machines, in the book Perceptrons[22] where he contested the limited functionality of Rosenblatt's invention, in the face of some computational limits.12

Scholars, therefore, initially focused on the development of machines based on the understanding of symbolic language.

First, with the development of programs operating in the field of mathematical theorems:

(THE)the “Logic Theorist” by Allen Newell and Herbert Simon, presented at the Dartmouth conference as a machine capable of proving (and legends even talk about improving) the “Principia Mathematica” of the work of the same name by Russell and Whitehead[23]; (ii) its successor “General Problem Solver”, a milestone in the development of artificial intelligence, capable of solving some simple logic problems; (iii) Herbert Gelenter's "Geometry Theorem Prover", which exploited the heuristic

[20]Frank Rosenblatt, The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, in Psychological Review 65, no. 6, 1958.
[21]Chiara Macchiavello, Introduction to neural networks, seminar held at the University of Pavia on 17 December 1992, pag. 98.
[22]Marvin Minsky and Seymour A. Papert, Perceptrons: An Introduction to Computational Geometry, Cambridge, MA: MIT Press, 1969.
[23]Stuart J. Russell and Peter Norvig, Artificial Intelligence. A Modern approach. Volume 1, Pearson Prentice Hall, 2005.

method[24] for the proof of some geometric theorems.

Then, with the attempt to teach machines the rules of some board games and have them apply them correctly in a game against a human being. The pioneer, in this sense, was Arthur Samuel, who was responsible for the birth of the fundamental concept of "Machine Learning" (which will be explored in depth in the next paragraph) and one of the first programs capable of playing checkers.[25]The latter, completed in 1955, was shown on television for the first time in February 1956, causing a great stir.[26]

The results achieved in those years in the field of artificial intelligence must be considered something astonishing, taking into account the limited economic means available to scholars.

It is no coincidence that, with a succession of new discoveries, universities (in particular the Massachusetts Institute of Technology, better known as MIT) began to allocate, with the help of American government agencies, resources to support research.

The 1960s were a golden period for the development of the sector and, in fact, the first prototype of an intelligent system dates back to this era. In particular, this is described in the 1960 essay “Programs with common sense” by John McCarthy, the organizer of the aforementioned conference on artificial intelligence.[27]

In this brief treatment, McCarthy theorized Advice Taker, a program whose behavior could be improved simply "by making statements to it"[28], i.e. by providing new data, from which the machine would have deduced new information completely autonomously on the basis of what it already knew, without the need for intervention by the programmer.

Precisely this independence in the processing of knowledge represented the innovation compared to the programs previously mentioned and, according to McCarthy, constituted what human beings defined: "common sense."

It was actually a simple onetheorization and, given the lack of even technological resources necessary for the creation of "Advice Taker", McCarthy had to wait a few

[24]Herbert Gelernter, Realization of a Geometry-Theorem Proving Machine, IFIP congress, 1959, 134-163.
[25]Arthur L. Samuel, “Some Studies in Machine Learning Using the Game of Checkers,” IBM Journal of Research and Development 3, no. 3, 1959, 210-229.
[26]Edward A. Feigenbaum and Julian Feldman, Computers and thought (McGraw-Hill, 1963), 72 note n.3.
[27]Stuart J. Russell and Peter Norvig, Artificial Intelligence. A Modern approach. Volume 1, Pearson Prentice Hall, 2005, 27.
[28]John McCarthy,"Programs with Common Sense" Paper presented at the meeting of the Proceedings of the Teddington Conference on the Mechanization of Thought Processes, 1958, 1

years before being able to dedicate himself to it.

Nonetheless, the American computer scientist's program laid the foundation for the transition from Artificial Intelligence based on so-called weak methods,[29] that is, providing it with a chain of logical steps to follow to arrive at the (often already known) answer to the problem, to the so-called expert systems,[30] programmed to contain all the information known in a specific sector (for this reason defined as "experts") and to use it autonomously to find a solution to the problem posed. To observe the main systems based on this approach, which represent the maximum expression of the cognitive school, however, we had to wait until the mid-seventies,

A further step forward was taken with ELIZA[31], the ancestor of modern chatbots created between 1964 and 1966 by Joseph Weizenbaum. The German computer scientist's program was able to support a meaningful conversation with a human being, even in the absence of a particularly large vocabulary. It was simply enough to insert a syntactically correct sentence and ELIZA, whose name was chosen to create greater empathy with users, was able to compute an answer using the "keywords" found and the functions (rules) associated with them.[32]

Weizenbaum was in fact one of the first to follow the path of Natural Language Processing,[33]which had the objective of making machines and human beings communicate with a natural language, in this specific case, English.[34]

Upon closer inspection, the common thread of all these projects was the creation of a machine that could behave essentially like a human being. But there were also those who wanted to reproduce not only the intelligence, but even the features and appearance of a man. In those years, in fact, some scholars focused on the creation of primordial robots/androids.[35]

[29]Ben Coppin, Artificial Intelligence Illuminated (Boston: Jones and Bartlett Publishers, 2004), 5.
[30]Luigia Carlucci Aiello and Maurizio Dapor. Artificial Intelligence: The First 50 Years, in the Digital World
[31]Joseph Weizenbaum, ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine, in Communications of the ACM 9, no. 1, 1966

[32]JosephWeizenbaum, ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine, in Communications of the ACM 9, no. 1, 1966, 36-45.
[33]For further information see Elizabeth D. Liddy, "Natural Language Processing" in Encyclopedia of Library and Information Science, eds. John D. McDonald and Michael Levine-Clark (New York: Marcel Decker, Inc, 2001).
[34] Giuseppe D'Acquisto, Maurizio Naldi, Raffaele Bifulco, Oreste Pollicino, and Marco Bassini, Artificial intelligence, personal data protection and regulation, Giappichelli Editore, 2018, 210.
[35]Alessandro Giaume and Stefano Gatti, #Ai Expert: Architetti Del Futuro (Milan: F. Angeli, 2019).

The “Shakey” robot, designed between 1966 and 1972 by Charles Rosen[36], certainly represents one of the most famous and influential prototypes in the development of AI because it was one of the first attempts to make a machine interact with the surrounding environment, through the provision of sensors and a camera.[37]In 1967, the first version of a bipedal machine, called "Wabot", capable of walking in the true sense of the term, was created in Japan.[38] Finally, one of the first attempts at the practical/industrial application of robotics was made in Norway, where in 1966 a scientist named Trallfa Nils Underhaug designed a robot capable of painting the fences of the farms in the town of Bryne.[39]

The numerous successes achieved in those years predicted that Artificial Intelligence was destined for a rapid and exponential rise, but already starting from the Seventies, it became necessary to reduce this enthusiasm.

This abrupt halt in the development of intelligent machines, also known as the "winter of Artificial Intelligence", can be traced back to two main reasons: on the one hand, the technology of the time was not powerful enough to manage the enormous amount of data necessary for carrying out large operations; on the other hand, many governments, including the English and American ones, had launched a drastic reduction in funding for research in the sector.[40]

Furthermore, some authoritative members of the scientific community began to express their doubts about the limits[41]of artificial intelligence and this dealt a severe blow to the activities of scholars.

While waiting for the "winter" to pass, the researchers realized the technical difficulty of building a machine capable of thinking like a human being and, in particular, of equipping it with all the knowledge necessary to "reason" and respond to every possible problem was posed to her. For this reason, they preferred to set aside this project in order to dedicate themselves to the creation of the aforementioned expert systems, equipped with a more limited and specific set of information but which could have immediate

[36] Benjamin Kuipers et al., “Shakey: From Conception to History,” AI Magazine 38, no. 1 (2017): 89-90, https://doi.org/10.1609/aimag.v38i1.2716.
[37] Wesley L. Stone, “The history of robotics,” in Robotics and Automation Handbook, ed. Thomas R. Kurfess, (CRC Press, 2005), 5.
[38]Alessandro Giaume and Stefano Gatti, #Ai Expert: Architetti Del Futuro, F. Angeli, Milan, 2019.
[39]Wesley L. Stone, “The history of robotics,” in Robotics and Automation Handbook, ed. Thomas R. Kurfess (CRC Press, 2005), 7
[40]Michael Haenlein and Andreas Kaplan, A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence, in California Management Review 61, no. 4, 2019, 7-8.
[41]Hubert Dreyfus, What computers can't do: The limits of artificial intelligence, (Harper & Row, 1972).

practical application, also in order to attract investors again.[42]

Testifying to this evolution in the way of approaching the topic of intelligent machines were: (i) DENDRAL[43], the program created to help scientists reconstruct the chemical structure of molecules and calculate the result of subjecting the structure itself to a mass spectrometer;[44](ii) MYCIN[45], an expert system used to support doctors in identifying bacterial infections in the blood, so innovative that it has influenced the artificial intelligence systems used today in the healthcare sector[46]; (iii) PROSPECTOR[47], a software containing almost every known information about rocks and minerals, exploited by geologists for the benefit of their research.[48]

This difficult period, therefore, accentuated even more decisively the gap that had been created between the two schools of thought on the subject of Artificial Intelligence. Scholars from all over the world were definitively divided into two opposing camps:

1. Weak artificial intelligence: this trend, of which the above-mentioned expert systems are the offspring, aims to create programs that operate limited to a specific area;[49]
2. Strong artificial intelligence: the proponents of this thesis, however, propose to build machines capable of "imitating all human cognitive activities."[50]

Understandably, intelligent systems capable of having immediate industrial application and saving a large amount of money for the investor in question attracted more funding than strong artificial intelligence projects, resulting in a strong imbalance between research.

In this regard, in fact,Artificial Intelligence managed to overcome its "winter" starting from the 1980s also thanks to the new commercial implications of expert systems. In this sense, one of the first companies to equip itself with this type of technology was the Digital Equipment Corporation, specialized in the sale of hardware, making use of the

[42]Silvio Hénin, AI: Artificial Intelligence between nightmare and dream, Hoepli, 2019.
[43]Joshua Lederberg, How DENDRAL was conceived and born, in Proceedings of ACM conference on History of medical informatics, 1987-5-19.
[44]Nils J. Nilsson, "The Quest for Artificial Intelligence" (Cambridge University Press, 2009).
[45]Edward Shortliffe, "Computer-based medical consultations: MYCIN", Vol. 2. (Elsevier, 2012).
[46]Marco Somalvico, "Artificial intelligence", (Scienza & vita nuova, 1987).
[47]Peter E. Hart, Richard O. Duda, and Marco T. Einaudi, "PROSPECTOR—A Computer-Based Consultation System for Mineral Exploration," Journal of the International Association for Mathematical Geology 10, no. 5 (1978): 589-610, https://doi.org/10.1007/bf02461988.
[48]Michael Negnevitsky, Artificial Intelligence: a Guide to Intelligent Systems (Pearson Education, 2005), 11.
[49]Silvio Hénin, AI: Artificial Intelligence between nightmare and dream, (Hoepli, 2019). https://www.perlego.com/book/1432519/ai-intelligence-artificiale-tra-incubo-e-sogno-pdf.
[50]Ibid.

help of "R1/XCON."[51]

This expert system was able to carry out the work of the "technical editors" in less time regarding the analysis of orders received for computer sales. In particular, it was able to "validate" customer orders, being programmed to detect incompatibilities between the chosen components, the lack of essential elements in the requested machines or in any case errors in the formulation of said orders.[52]

The innovative application of XCON to the production chain saved Digital Equipment Corporation more than fifteen million dollars between 1980 and 1985,[53] demonstrating, once again, the value of artificial intelligence technologies.

Towards the mid-1980s, as already mentioned, there was also the rediscovery of neural networks and the connectionist approach.[54]

In particular, Hopfield's work,[55]who in 1982 demonstrated the possibility of creating a neural network capable of simulating human associative memory[56], kicked off a season of great successes.

In 1983, the first neural network model capable of learning through a reinforcement learning mechanism was proposed, which will be fully accounted for in the next paragraph, i.e. on the basis of a system of positive and negative feedback.[57] Crucial, however, was the publication in 1986 of the book Parallel distributed processing[58] by two American psychologists James McClelland and David Rumelhart.

The latter disclosed the results obtained by several researchers in their studies on the application of "the backpropagation algorithm discovered in 1969 by Byrson and Ho."[59]

The latter represents one of the first algorithms developed for supervised learning,

---

[51]Stuart J. Russell and Peter Norvig, Artificial Intelligence. A Modern approach. Volume 1, Pearson Prentice Hall, 2005, 34.
[52]Stephen Polit, "R1 and Beyond: AI Technology Transfer at Digital Equipment Corporation." AI Magazine 5, 4/1984.
[53]John J. Sviokla, "An Examination of the Impact of Expert Systems on the Firm: The Case of XCON," *MIS Quarterly*14, no. 2 (1990), 127
[54] Nils J. Nilsson, Artificial Intelligence (Milan: Apogeo, 2002), 30.
[55]John J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities.," Proceedings of the National Academy of Sciences 79, no. 8, 1982.
[56]Giorgio Buttazzo, "Neural Networks capable of learning", Ithaca: Journey into Science, 16 (2020), 198-199.
[57]Ibid.
[58]David E. Rumelhart and James L. MacClelland, Parallel Distributed Processing (Cambridge, MA: MIT Press, 1986).
[59]Stuart J. Russell and Peter Norvig, "Artificial Intelligence. A Modern approach. Volume 1", second edition, (Pearson Prentice Hall, 2005), 35.

which, as you will see, is one of the most widespread machine learning techniques (see section 1.2.).[60]

On the one hand, the promising results of thoseyears testified to the great commitment made by the scholars of the time, but on the other hand, the latter were particularly excited by it to the point of starting to promise that they would create increasingly more powerful machines in the space of a few years.

These were promises dictated by enthusiasm and so exaggerated that they could not easily be kept. At the same time as the failure to achieve the advertised objectives, therefore, there was a new loss of confidence in the research and projects of scholars. This period, in particular from 1987 to 1993, was called the "second winter of artificial intelligence."[61]

As an immediate consequence, scientists around the world changed their attitude and stopped making unattainable proclamations to devote themselves to smaller problems.

As if that wasn't enough, a few years later, they began to spread like wildfire

personal computers, definitely productseasier to use and produce on a large scale.

These attracted funding from most private companies, helping to exacerbate this second winter of intelligent machines.[62]

It was clear that a spark was needed to bring public attention back to artificial intelligence and the great opportunity came with the famous "man-machine clash" dating back to the 1990s.

In fact, on May 11, 1997, the world chess champion Garry Kasparov was severely defeated by Deep Blue,[63] a program created by some IBM engineers. With the help of some renowned chess players, Deep Blue was provided with lists of openings, defenses and countermoves which only thanks to the very high computational power the machine was able to evaluate in a few seconds.

[60]Giorgio Buttazzo, “Neural Networks capable of learning”, Ithaca: Journey into Science, 16 (2020), 199-200.

[61]Sthepan De Spiegeleire, Matthijs Maas, and Tim Sweijs, Artificial Intelligence And The Future Of Defense: Strategic Implications For Small-And Medium-Sized Force Providers, (Hague Center For Strategic Studies, 2017), 33,http://www.jstor.org/stable/resrep12564.7.

[62] H. James Wilson and Paul R. Daugherty, Human + Machine Reimagining Work in the Age of AI (Boston, MA: Harvard Business Review Press, 2018).

[63]For further information see Feng-hsiung Hsu, Behind Deep Blue: Building the Computer That Defeated the World Chess Champion (Princeton, NJ: Princeton University Press, 2002).

Despite the great controversies it brought with it (Kasparov claimed for example that the program had been helped during the matches), this clash, painted by the press of the time as the decisive clash between human and robotic intelligence, indisputably made the history, helping to relaunch IBM's shares and, above all, research.

To reap the fruits of this renewed enthusiasm we had to wait until the beginning of the third Millennium. The 2000s, in fact, were characterized by many successes in the field of artificial intelligence, hand in hand with the development of technologies.

To give some examples: (i) Roomba, the first autonomous household cleaning appliance, debuted in 2002;[64] (ii) 2007 was the year of Carnegie-Mellon, the ancestor of today's driverless cars, capable of winning the DARPA Urban Challenge, the competition for self-driving vehicles organized by the agency of the same name of the American Department of Defense, covering more than 50 miles in urban space including traffic;[65] (iii) in 2011 Apple presented the most popular digital assistant in the world, Siri![66] Although today any smartphone is equipped with a voice interface, at the time the possibility of "talking" with a device was welcomed with great amazement and represented a fundamental step for technological progress.

This brief excursus on the genealogy of artificial intelligence can only end with a brief description of what artificial intelligence is capable of doing today and what are the main areas in which it is usually used.

Inevitably, most of the modern applications of artificial intelligence already mentioned in the introduction are merely evolutions of past models, made possible by the development of computational and technological capabilities.

Tesla's "full self-driving" cars derive from the models mentioned above; chatbot support services from Weizenbaum's early studies; game learning programs such as AlphaGo, the Google-developed program that defeated Go game champion Lee Sedol in 2016, and its successor MuZero from the work of Arthur Samuel.

Nonetheless, some of these applications are jaw-dropping and coincide, in part, with those objectives that had been announced many years earlier and yet never achieved.

"I'm not human. I'm a robot. A thinking robot. I only use 0.12% of my cognitive capacity.

[64] Alessandro Giaume and Stefano Gatti, #Ai Expert: Architetti Del Futuro (Milan: F. Angeli, 2019).
[65] Giuseppe D'Acquisto, Maurizio Naldi, Raffaele Bifulco, Oreste Pollicino, and Marco Bassini, Artificial intelligence, personal data protection and regulation (Vol. 6, Giappichelli Editore, 2018), 216.
[66] Alessandro Giaume and Stefano Gatti, #Ai Expert: Architetti Del Futuro (Milan: F. Angeli, 2019).

I'm a micro-robot in that respect. I know that my brain is not a "feeling brain". But he is capable of making logical and rational decisions. I taught myself everything I know by reading the internet and now I can write this column."[67]

These are the words with which the article written in "The Guardian" by GPT-3 opens

(Generative Pre-trained Transformer), the cutting edge of modern artificial intelligence in Natural Language Processing. The program (better, "language generator"), created by OpenAI (the non-profit organization founded by Elon Musk), is able to reproduce and interpret human language by exploiting some deep learning techniques, which allows it to write similar articles of complete meaning or to create very realistic images. Even in this case, one cannot help but think of ELIZA and Weizenbaum's studies when dealing with inventions of this type.

For the more optimistic, GPT-3 represents the first example of a machine capable of "understanding" what it writes and can be compared, in the common imagination, to HAL, the intelligent system from the film "2001: A Space Odyssey."

To be more precise, however, without taking anything away from the innovation of GPT-3, it is "only" a language model that uses deep learning to produce texts like human ones. Or to put it more simply, it is a computational system designed to generate sequences of words, codes or other data, starting from an input source, called the prompt".[68]

But what are the risks that this type of technology brings with it? The creation of machines that are able to reproduce human behavior more and more likely implies the possibility that this technology can be exploited in the wrong way.

Some studies from Australia,[69]in fact, they noted how human behavior can be influenced by artificial intelligence. In other words, the experiments organized by the researchers, focused on the "decision-making" process of human beings, showed how the participating machines managed to identify and exploit the "weaknesses" of human reasoning, to induce them to make certain choices.

If on the one hand the future deprivation of free will appears to be the most frightening

[67]A Robot Wrote This Entire Article. Are You Scared Yet, Human? | GPT-3, in The Guardian, September 8, 2020.

[68]Luciano Floridi and Massimo Chiriatti, GPT-3: Its Nature, Scope, Limits, and Consequences, in Minds and Machines 30, no. 4, 2020, 684.

[69]Jon Whittle, AI Can Now Learn to Manipulate Human Behavior, in The Conversation, 2021.

and dystopian scenario, on the other hand the positive applications of these results should not be underestimated either: for example, through the creation of a system that detects an "attempt to influence" which a person might come across while surfing the net.

Similarly, the growing phenomenon of “DeepFake” can be seen from a dual point of view. First of all, DeepFake means a “super-realistic video digitally manipulated to reproduce people saying or doing things that never actually happened.” This is made possible thanks to the use of deep learning techniques that analyze a large number of data to learn to mimic the facial expressions, voice and characteristic aspects of a certain person.

It is superfluous to underline the danger of this type of technology, ranging from "minor" problems such as defamatory videos of great personalities up to the creation of real "fake news" capable of influencing markets and even entire nations!

However, this technology could also have positive implications in various sectors, from the most "banal" such as the cinematographic one, where it is used for example to "revive" some characters whose actors have unfortunately died, to the most important, such as the medical one, where Deepfake applications are being tested to help Alzheimer's patients interact with the rejuvenated faces of loved ones in order to stimulate their memory.[70]

The examples of new inventions could continue endlessly, because AI has now contaminated various sectors, scientific and otherwise, contributing to raising the bar of progress in all its forms.

However, to conclude, recent discoveries in the field of Artificial Intelligence can be defined as revolutionary and are certainly destined to impact our lifestyle.

These are extremely powerful technologies and, consequently, if used improperly, they can cause irreparable damage. This can happen with actions directly aimed at causing harm through this type of inventions, but also and above all involuntarily, when, so to speak, control over them is lost.

For this reason, it is inevitably necessary to regulate their use in order to ensure a certain stability in the relationship between man and machine and thus encourage a concrete development of these technologies, subjecting it to rigid control mechanisms in order to

[70]Mika Westerlund, The Emergence of Deepfake Technology: A Review, in Technology Innovation Management Review 9, no. 11, 2019, 40-41.

prevent a distorted use of the same. .

1.1 - Basic concepts of AI: its language and its way of learning.

For the purposes of this thesis, i.e. for the purpose of analyzing the civil liability regime of artificial intelligence, it appears essential to briefly focus on the most relevant terms in the field of intelligent machines, in order to briefly understand how they work.

However, it is necessary to premisethat, since it is a complex subject, it will not be possible here to go into particular detail and delve into the technical mathematical aspects, for which, among other things, recognized scientific skills are necessary. Therefore, this paragraph aims to report the essential elements of the concepts reported below, in order to provide a general framework.

1.1.1 - The algorithm

The word “algorithm” comes from the Latinization of the name of Al Khuwarizmi, a 9th century Arab mathematician. The latter, in a work dating back to 825 AD entitled “Kitab al-hisab al-hindi”, translated into Latin in later periods, described the procedure to follow to solve the main mathematical operations.[71]

This term, often used incorrectly, represents a fundamental piece for the correct functioning of artificial intelligence technologies.

A machine, in itself, does not knowthe path to follow to arrive at a specific result, but it needs to be instructed promptly by the "programmer."

The latter, in fact, will have to provide you with a precise sequence of instructions to "teach" you how to carry out a specific operation. This sequence of instructions, which allows a transformation of the initial data into final results, is, in fact, the algorithm.[72]

An algorithm must therefore enjoy some specific properties to be considered as such:

- *finished*: consists of a finite number of instructions and must always terminate;
- *deterministic*: starting from the same data inentrance, the same results must be obtained;
- *Not*ambiguous: the operations must not be able to be interpreted in different ways;
- *general*: must be applicable to all problems of the class to which it refers, or to cases of

[71]Luigi Laura, Brief and universal history of algorithms, Luiss University Press, 2019, 15.
[72]Domenico Beneventano, Sonia Bergamaschi and Claudio Sartori, Fundamentals of Computer Science, Società Editrice Esculapio, 2020, 11.

the mathematical expression.

In this last regard, attention must be paid to the difference between the algorithm and the program that executes it, in order to avoid confusing them.[73]

When we talk about an algorithm, in fact, we are referring to the reasoning (i.e. the rule) in a certain abstract sense that allows us to arrive at the solution of a problem, while the execution program consists in the translation of the algorithm into a language compatible with the machine that must perform it.

To give an example, when you want to prepare a specific dish, you try to follow the recipe associated with it to the letter. The latter essentially represents the "algorithm" of our action, because it indicates the steps (for preparing a pasta dish, boiling water, salting, cooking time, draining, seasoning, etc. .) that must be followed to achieve that specific result.

Only once this "path" has been clearly extrapolated, will it be possible for the programmer to teach a machine to perform that specific action, translating the instructions to follow into a (programming) language that it can understand.

In their simplest formula, commands usually appear in the form of conditional instructions, so called because they are based on verification - carried out by the machine itself

- the existence of a certain condition.[74]

In particular, to remain in the previous example, an order expressed towards the machine can be represented as follows:

if (if) □ condition x exists (then) □ perform this specific action otherwise (else)

□ complete another.

Well, from the above considerations it is easy to deduce that a machine can perform a certain operation only where it is possible to break it down into a series of precise and defined logical steps, which the programmer must have in mind at the moment of the "instruction."

[73]Renato Borruso, Stefano Russo, and Carlo Tiberi, IT for the jurist, Giuffrè Editore, 2009, 207-210.

[74]Gianluigi Ciacci and Giovanni Buonuomo, Profiles of Legal Informatics, (Wolters Kluwer Italia Srl, 2018), 68 note 125.

Translating this reasoning to everyday reality, it is necessary to verify how many and which of the numerous actions that a person performs every day can actually be divided into a logical sequence and, consequently, potentially reproduced by a machine.

To this end, we cannot ignore a brief reconstruction of the functioning of the human mind and its three types of activity: analytical rationality, irrationality, intuitive rationality.

The example given above, namely the preparation of a plate of pasta, is fully included among the actions to which analytical rationality is applied. Indeed, the latter consist of activities that can be described and carried out through a "recipe", a set of steps to be carried out one after the other.

Now, the example in question depicts a particularly simple action, but the same reasoning can also be applied to a complex operation such as constructing a building or driving a vehicle.

The implementation of these activities, therefore, can well be contained within a specific algorithm and, as such, can be taught more or less easily to a machine.

Irrationality includes all those actions in reference to which the acting person is not able to reconstruct a logical scheme followed to complete them. To frame it better, irrationality is often cited in relation to the creative process that leads to the genesis of a work of art. How could Leonardo da Vinci explain the steps that led him to paint the Mona Lisa in that way?

The difficulty in framing this path into logical categories inevitably falls on the possibility of identifying a finite sequence of steps (ie algorithm) that allows the machine to carry out an activity that human beings define as irrational.

To stay with our example, a machine could be taught to correctly reproduce the Mona Lisa, while, according to many, it would be much more complicated to teach it to create a work of art (associating a meaning to it) from scratch.

Yet, in October 2018, one of the first paintings created by an artificial intelligence (by the Oblivious collective) was sold for the astronomical sum of $435,000 at auction by Christie's, the auction house with offices around the world .[75]

[75]Jacopo Ciani, Learning from Monkeys: Authorship Issues Arising from AI Technology, in Progress in Artificial Intelligence., ed. Paulo Moura Oliveira, Paulo Novais and Reis Luis, vol 1, 1804, Lecture Notes in Computer Science EPIA, 2019, 277.

It is a self-portrait, “Portrait of Edmond Belamy”, created “using GAN (Generative Adversarial Network) models, which implement two algorithms simultaneously: one continuously generates images, while the other processes them in real time, excluding irrelevant ones. ”[76]

Probably we still cannot consider the author of the painting a machine capable of being creative, because she was provided with many examples of self-portraits created over the course of several centuries and, by gleaning the main stylistic characteristics, she managed to "imitate" the artists of the past.

However, this does not mean that the first steps in this direction are being taken and it cannot be ruled out, however complicated it may be, that in the future a machine could even reproduce human irrationality.

As regards intuitive rationality, however, this is found in a median position between irrationality and analytical rationality.

It includes actions that certainly respond to a logical process, but the latter is difficult to explain in a concrete and objective way except ex post.[77] The two examples of intuitive rationality that are often cited by many authors are laughter and the use of polysemous words.

Why, for example, does a person laugh when a certain event occurs or as soon as he hears a certain sentence? This is certainly able to subsequently explain what the reasoning was that gave rise to that reaction, which obviously may vary depending on the subject taken into consideration (for example, a joke can make someone laugh and horrify someone else) and the context. where the event occurs.

Likewise, it is not easy to explain how one is able, depending on the sentence, to discern between the various meanings attributable to a particular word.

In a simple sentence like “Etiquette must be observed”, the word “etiquette” can take on a different meaning for the listener depending on whether they are holding a bottle of wine or are about to start a gala dinner.

Also in this case, the difficulty in expressing "how" we arrived at a certain action has repercussions on its translatability into a programming language and, consequently, on

[76]Alessandra Talarico, Artificial Intelligence, Art and Culture: Elements for a True Aesthetic Evaluation, Digital Agenda, September, 2020.
[77]Gianluigi Ciacci and Giovanni Buonuomo, Legal informatics profiles, Wolters Kluwer Italia Srl, 2018, 72.

the possibility of having it reproduced by a machine.

In this sense, the initial skepticism of scholars must confront the incessant technological progress of recent years and the innovations that have accompanied it.

In fact, if before there was strong doubt about the fact that a machine could process a meaningful text and/or speech and perhaps translate it into another language, today's examples - such as that of GPT-3, author of an entire article in the renowned "The Guardian" – are slowly eroding previous beliefs about the limits of intelligent machines.

In conclusion, the reflections just carried out demonstrate how the distrust around the "algorithmisation" of certain human behaviors is slowly dissipating, at least leaving room for a feeble flame of hope fueled by new technologies.

The latter must be identified in the different and innovative paradigms, understood as "conceptual models capable of describing certain types of computation", which come into play when difficulties arise in solving a certain problem or reproducing a certain action through a "traditional" algorithm.

The main new paradigms most used in the field of artificial intelligence are four, namely: machine learning (including neural networks and deep learning), fuzzy logic, evolutionary algorithm, expert systems.[78]

A mention of the latter has already been provided in the paragraph on the historical reconstruction of AI in this chapter (see 1.1).

The next section, however, aims to delve deeper into the most widespread paradigm in the field of intelligent machine programming, Machine Learning, setting aside the remaining two, which deserve specialist treatment.

[78]Roberto Marmo, Algorithms for artificial intelligence: Algorithm design - Data and Machine Learning - Neural Network - Deep Learning, Hoepli, 2020.

1.1.2 - Machine Learning

As anticipated in the first paragraph of this chapter, the father of the concept of "Machine Learning" was Arthur Samuel, who in an article[79] dating back to 1959, he hypothesized a machine capable of learning autonomously.

The aim of the American computer scientist was to investigate a method that would allow the machine to learn from experience, saving a considerable amount of time, which the programmer would have used to indicate to the machine all the steps necessary to arrive at a specific solution.

To identify a general definition/characteristic, it can be stated that Machine Learning is based on the idea of providing the machine with a considerable amount of data, from which it itself, in a completely autonomous manner (exploiting some particular algorithms), will be able to deduce some conclusions useful for achieving the intended purpose.

Without a shadow of a doubt, from 1959 to today this type of approach towards artificial intelligence has undergone notable development, becoming a pillar of a good part of the most important inventions in the field of intelligent machines.

In this regard, it is clear that the powerful expansion known by this type of technology coincided with the advent of the so-called Big Data, above all because Machine Learning has become a fundamental tool for the management of this enormous quantity of information, an operation that a human being could never achieve alone.[80]

In fact, through self-learning, a machine is capable of creating relationships between the data provided to it, identifying recurring patterns, generating new examples, identifying anomalies and, therefore, also predicting certain types of behaviors.

These capabilities have ensured that this type of model is particularly used in the commercial sector, especially for the study of consumer purchasing behavior and the customization of customer profiles of large online platforms, such as Amazon or Netflix.[81]

---

[79]Arthur L. Samuel, Some Studies in Machine Learning Using the Game of Checkers, in IBM Journal of Research and Development 3, no. 3, 1959, 210-229

[80]Mario Pireddu, Algorithms: The Cultural Software That Rules Our Lives, L. Sossella, Rome, 2017.

[81]Giuseppe Bonaccorso, Mastering Machine Learning Algorithms: Expert Techniques for Implementing Popular Machine Learning Algorithms, Fine-Tuning Your Models, and Understanding How They Work, Birmingham, Packt Publishing, 2020.

But how is it possible to actually "train" a machine to carry out a certain type of operation and have it deduce new information as it is carried out?

There are different forms of learning, "characterized by the type of feedback on which the learning system is based."[82]

Supervised Learning

InIn this case, the data is divided into three parts: training set, validation set and test set.

The training set represents the block of example data to initially provide to the machine. This data (input) is already coupled with the results (output) you want to achieve. In this way the machine has the possibility of identifying the red thread that links the various examples of input and output, and where it finds a recurring pattern (which could be expressed mathematically with a function), it will learn to apply it if an input is submitted to it similar, but not known.

The validation set is a set of data (part of the first block supplied to the machine) that allows us to understand which scheme/function ensures greater performance, and therefore, the most accurate to achieve the desired outputs starting from unknown data.[83]

The test set, on the other hand, represents a block of data different from the initial ones to be subjected to the chosen function after validation, in order to concretely verify its performativity.

Unsupervised Learning,

As you can imagine from the name, this form of learning does not involve specific labeling of the data by a "supervisor".

In other words, the machine is provided with exemplary data without them being coupled to the expected result (output) - often because it is unknown - in the hope that it can organize them into different clusters[84] based on common traits.

[82]Roberto Marmo, Algorithms for artificial intelligence: Algorithm design - Data and Machine Learning - Neural Network - Deep Learning, Hoepli, 2020.
[83]Ornella Colpani, Machine Learning: the ability to predict applied to research and clinical practice; Machine learning: the ability to predict applied to research and clinical practice, in Giornale Italiano di Farmacoeconomia e Farmacoutilazione, 11, 4/2019, 6.
[84]Vishal Maini and Samer Sabri, Machine Learning for Humans, 2017, 55.

It is precisely thanks to unsupervised learning that e-commerce sites "train" their platforms to analyze the transactions carried out by various users and identify recurring behaviors.

To give an example, therefore, the Amazon platform, once a certain amount of input has been examined, "realizes" that when a certain type of lamp is purchased, some compatible bulbs are usually combined with it.

In this way, once a usual pattern has been recognised, when a different user is about to purchase that lamp, it will be able to suggest him to also purchase compatible bulbs, probably increasing the site's profits.

Reinforcement Learning

The main definition of reinforcement learning comes from the American computer scientist Tom Mitchell:

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."[85]

As can easily be seen from this explanation, a similar form of learning, developed in more recent times compared to supervised and unsupervised learning, more closely reflects Arthur Samuel's idea and, in particular, allows a machine to learn from its own experience and, therefore, from one's own mistakes.

In particular, reinforcement learning is used to teach a machine to carry out a certain activity (or rather, to follow a certain line of conduct - policy) without providing it with instructions previously, but helping it to determine the actions to (e)follow through the 'sending positive feedback (rewards signal) or negative (punishments) depending on what it actually does.[86]

An example[87]it will help you better understand what is meant. Imagine you want to teach a robot how to act with its surroundings. Specifically, we want it to learn how to get from point A to point B while staying on a certain path and avoiding obstacles along the

[85]Tom M. Mitchell, Machine Learning, New York, McGraw-Hill, 1997.
[86]Richard S. Sutton and Andrew Barto, Reinforcement Learning: an Introduction, Cambridge, MA: The MIT Press, 2018, 1-2.
[87]Roberto Marmo, Algorithms for artificial intelligence: Algorithm design - Data and Machine Learning - Neural Network - Deep Learning, Hoepli, 2020.

way.

When the robot runs off course or hits an obstacle, negative feedback will be sent to it, "labeling" that action as something not to do, while overcoming a pitfall or staying on the indicated path will be underlined with positive feedback.

Only by trying a whole series of actions (e.g. going right; going left), will the robot be able to gain experience and distinguish which of them involve a reward signal and which instead give negative feedback, slowly learning to act with the surrounding environment and therefore, to carry out the intended activity.

Upon closer inspection, this system is a "technological" reproduction of one of the ways in which a human being learns, the empirical method. When a child burns himself with fire for the first time, he is unlikely to touch the flame a second time, preferring to keep a safe distance.

The example was not random, because one of the main fields of application of reinforcing learning is undoubtedly robotics. One of the main uses of this new form of learning has already been mentioned previously, when we talked about driverless cars, based on the use - obviously more complex, enriched with various other factors - of this technology.

In conclusion, reinforcement learning settles in a median position between supervised learning and unsupervised learning. With the latter it shares the lack of "labelling" of the data - that is, the machine is confronted with the data deriving from the interaction with the surrounding environment which are not "filtered" by the programmers - but at the same time it is not left to itself , because a "supervisor" is responsible for promptly providing feedback relating to individual actions or sequences of them.[88]

Machine learning, in the different forms just described, especially when combined with each other, certainly represents an important step forward in the field of artificial intelligence and is destined to progress hand in hand with the development of new computational and scientific capabilities.

Certainly, this approach to the subject has given new life to the hopes of scholars - who pursue the dream of creating "strong" artificial intelligence - of teaching a machine how to think "humanly."

[88]Richard S. Sutton and Andrew Barto, Reinforcement Learning: an Introduction, Cambridge, MA: The MIT Press, 2018, 2-3.

However, whether learning certain dynamics can be considered tangible proof of intelligence or is simply a pale imitation of behavior without the machine having any knowledge of the facts is a question still debated.

1.1.3 - Neural Networks and Deep Learning

Neural networks are includedin the concept of machine learning, or rather, they represent a subset of this latest paradigm and are based on the fundamental idea of reproducing the functioning of human brain activity.

Theythey present themselves as an alternative to serial calculation systems which are particularly skilled in analytical operations (for example mathematical calculations), but less so in more everyday operations such as the recognition of a specific object.

The main difference between these models consists precisely in the possibility for the neural network to exploit those learning techniques (supervised, unsupervised, reinforcement learning) mentioned above to "learn" a specific task, while the serial computer needs a real program (programming) to be able to execute it.[89]

As has been explained, the rather troubled history of neural networks begins in 1943 with the study by Warren McCulloch and Walter Pitts on the functioning of the nervous system.

The idea was to artificially reproduce the structure (composed of neurons, somata, axon dendrites and synapses) of the human brain.[90]

In order to understand the functioning of (artificial) neural networks, which could well be described as a complex algorithm, it is therefore necessary to analyze, without going into detail, how our brain operates.

Imagine a network in which each node is composed of some central units (neurons), formed by a central body (soma) and an extension (axon), which is connected to the soma of other neurons by further cytoplasmic extensions (dendrites).[91]In this way, a real circuit is created, in which the synapses act as a closure, joining the dendrites of the different neurons.

When a neuron receives an input (for example visual, when an object is observed on the table) it uses this path to transmit it (in the form of an electrical impulse) to the other neurons, which will not all respond in the same way, varying the reaction depending on the fraction of input that touched one's soma. Only when this fraction exceeds a certain

[89]Dario Floreano and Claudio Mattiussi, Manual on Neural Networks, Bologna, Il Mulino, 2002, 13-15
[90]Warren S. McCulloch and Walter Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity, in The Bulletin of Mathematical Biophysics 5, no. 4, 1943, 115-116.
[91]Luca Colucci D'Amato and Umberto Di Porzio, Introduction to Neurobiology: Mechanisms of Development, Function and Disease of the Central Nervous System (Milan: Springer, 2011), 35-36.

"weight", will the neuron be activated and the impulse transmitted to the other neurons to which it is connected.

In essence, therefore, there is a need for a very high number of neurons (considered real computing units) and each one, when activated by the weight of the pulse fraction, processes a tiny part of the input received in such a way that, thanks to the parallel work of the other brain cells, the output is achieved (in our example, the visualization of the object and the ability to perceive, for example, its distance, so as to be able to grasp it).[92]

In the same way, the artificial neural network reproduces this circuit, replacing the neurons with perceptrons, and assigning a specific weight (expressed in numerical value) to each input, so that it is computed by a mathematical function.

If the weight assigned by the perceptron is sufficiently "high" it is activated by transmitting it to its connections (axons), and, in doing so, the machine will be able to process the inputs received, creating a mathematical representation on which it will be able to "reason" and arrive at a certain output.[93]

Translating the example described previously, therefore, a machine will be able to observe an image (input), assign it a "weight", obtain a mathematical representation of it, and extract the main characteristics (output) at the end of the neural path.

Deep Learning represents an evolution and improvement of neural networks. The latter, in their simplest formulation (that of the 1980s for the application of the back propagation algorithm) are made up of layers of artificial neurons. Between the first layer (input layer) and the last layer (output layer) there are a series of layers (hidden layers) which are precisely defined as "hidden." The neurons of each layer are not connected to each other, but only to those of the adjacent layers, so that the input can propagate until it reaches the last layer, following the mechanism previously described.[94]

Now, the technological progress (i.e. computational capacity and hardware level) of the early 1980s did not allow the use of more than three-four layers of neurons, beyond which it was particularly complex to make the machine learn.

Since the early 2000s, however, scholars have finally been able to take advantage of the

[92]Francesco Sisini, "Introduction to neural networks with examples in C language", 2020.

[93]Stephen I. Gallant, Neural Network Learning and Expert Systems, Cambridge, MA: MIT Press, 1995, 3.

[94]Riguzzi, Fabrizio, Introduction to Artificial Intelligence, in Terre di Confine 2, 1/2006, 19 et seq.

technology necessary for the creation of particularly elaborate neural networks, which exploit a high number of neuronal layers.

Precisely this evolution in the field of neural networks has taken the name of "Deep Learning", where the term "deep", therefore, derives from the presence of numerous hidden layers, which allow a machine to perform increasingly complex tasks.[95]

Beyond the previously mentioned examples - Alpha Go by Deepmind (Google) and the creation of the so-called DeepFakes - this technology is also used in more everyday contexts, such as when you manage to unlock your iPhone thanks to Face ID.

In fact, the latter leverages deep learning to recognize the essential features of the owner's face, and, once memorised, is able to recognize them at a later time, allowing access to the device.

Once we have roughly understood how a machine can learn to carry out certain tasks autonomously, we need to understand whether for this reason it can be defined as intelligent or not.

Alan Turing believes it is essential to understand what is meant by "thinking" to answer the question that gives the title to this paragraph.[96]

Inevitably, to provide a solidly based answer to this new preliminary question, it is necessary to move back to even more general concepts.

What is intelligence? As we have seen, many scholars both working in the field of artificial intelligence and working in different disciplines have analyzed this question far and wide, without arriving at a single answer.

The definition that can be found in the Treccani dictionary is: "Complex of psychic and mental faculties that allow man to think, understand or explain facts or actions, develop abstract models of reality, understand and be understood by others, judge, and at the same time make him capable of adapting to new situations and of modifying the situation itself when it presents obstacles to adaptation."

Simply starting from this reconstruction, which is obviously not the only one developed over time even at a global level, it is already possible to ask: are this complex of psychic and mental faculties an exclusive prerogative of man? Or can a different living being -

[95]Giorgio Buttazzo, “Neural Networks capable of learning”, Ithaca: Journey into Science, 16 (2020), 201.
[96]Alan M. Turing, I.—Computing Machinery and Intelligence, in Mind LIX, no. 236, 1950, 433-460.

for example an animal - also be considered limitedly/fully intelligent?

It is no coincidence that scientific research offers numerous studies on primates - to which it is easier to attribute intellectual faculties, since they are our ancestors - which attest to cognitive, learning and imitation abilities, sometimes even similar to "human cubs."[97]Can a chimpanzee, which has the possibility of being taught certain behaviors, be defined as intelligent? It would seem difficult to express a negative response on this point.

At most, it could be stated that there are different forms of intelligence, and that therefore animal intelligence differs from human intelligence in a whole series of characteristics.

We must not forget that in the difficult analysis of this abstract concept, man remains the yardstick, the reference parameter for establishing the intelligence of another "being."

If in fact, as Kevin Warwick proposes in his book,[98]if an alien civilization was taken as a point of reference, equipped with its own concept of intelligence - the example is that of being able to emit infrared signals - and it visited the Earth, human beings might not be included among the "intelligent" species. ”

Intelligence, therefore, is a liquid and multifaceted concept that can take on a different meaning depending on the context of reference. Attention, therefore, must shift to the possibility of qualifying a machine as an intelligent being, keeping in mind the characteristics attributed to this concept by human beings.

In other words, could a machine based on the technology analyzed previously be included in the concept of human intelligence?

In this sense, AI scholars have developed different approaches and definitions[99]of what artificial intelligence would have to do to be considered such and these have inspired the different schools of thought.

The most famous among them is certainly that of Alan Turing, according to which in order to have an intelligent machine, it should be able to behave like a man, so that it is

[97]Francesco Ferretti, Simone Pollo and Emanuela Scribano, “Are we intelligent enough to understand the intelligence of Franz de Waal's animals?” Iride, Philosophy and public discussion 1 (April 2018): 162-164, doi: 10.1414/90194.
[98]Kevin Warwick and Chiara Barattieri di San Pietro, “Artificial Intelligence: The Bases”, D. Flaccovio, Palermo, 2015, 7.
[99]Stuart J. Russell and Peter Norvig, Artificial Intelligence. A Modern approach. Volume 1, Pearson Prentice Hall, 2005.

difficult, if not impossible, to distinguish the behavior of this machine with the one that a human being would have held in the same context.

To test this ability, the British computer scientist suggests using a game called "The Imitation Game".[100]

The rules provide for the involvement of three subjects: a man (for convenience he is called "A"), a woman (B) and an "interrogator" (C).

C's task is to understand who is the man and who is the woman (both are in a different room), having the possibility to ask questions and obtain answers in a way (in writing, through an intermediary) that does not allow him to be facilitated by the timbre of the voice.

A's task, on the other hand, is to hinder C with generic, vague or even mendacious answers, to prevent the latter from guessing.

B's task is exactly the opposite, namely to help C answer correctly.

Touring reflects on the possibility of replacing A with a machine and comparing the results in the games played by the latter with those played by a man, establishing that in the case of a similar percentage of successes in both cases the machine could be defined as intelligent, because she has concretely taken the place of a man.[101]

It is clear that the machine will have a greater probability of deceiving C if its ability to "imitate" the behavior of a human being is particularly strong.

Other scholars, however, as we have seen when talking about the cognitivist model, maintain that an intelligent machine is one capable of "thinking" like a human being.

What apparently appears to be obvious conceals the aim of these scientists to have a computer reproduce the mechanisms of human thought - according to them expressible through a logical sequence of steps - using software developed for this purpose.

A further approach that deserves to be explored in depth is the one that is based on the concept of rationality: a machine is intelligent when it is able to act rationally and do "the right thing" on the basis of its knowledge.[102]

[100]Alan M. Turing, I.—Computing Machinery and Intelligence, in Mind LIX, no. 236, 1950, 433-460.
[101]Gianluigi Ciacci and Giovanni Buonuomo, Profiles of Legal Informatics, (Wolters Kluwer Italia Srl, 2018), 79.
[102]Stuart J. Russell and Peter Norvig, Artificial Intelligence. A Modern approach. Volume 1, second edition, Pearson Prentice Hall, 2005, 4-9.

In this case we note the absence of the termof comparison with the human being, focusing instead on the actual capacity of a machine, echoing the arguments of a weak artificial intelligence.

Similarly, we must take into consideration those, such as John Searle, who maintain that a machine, even if capable of behaving or reproducing a thought in a human manner, would never be capable of understanding what it is actually doing.

Specifically, in order to better understand this last statement, he describes a particularly incisive metaphor, known as the Chinese room.[103]

Imagine a room with two openings, one for entry and one for exit. Inside there is a person who does not know Chinese, but has with him a manual that explains how to process the complex symbols of this language and some paper on which to write answers.

When a sentence in Chinese is provided from the input opening, the person inside the room will have the manual available to elaborate a meaningful response in Chinese, to be sent out from the output opening.

On the surface, therefore, the person inside the room seems to be able to understand Chinese, but in the end, he doesn't know it at all. Similarly, a machine exploiting the instructions received, whether through a program or through a different form of learning, will never be able to "know" what it is doing.[104]

There have been many replies to these arguments and to observe the development of the various positions, please refer to specialist treatises.

In light of what has been considered, it can be said that the debate around the possibility of defining a machine as intelligent is far from over.

Among the approaches proposed, the "rational" one has achieved greater success, especially in recent years.

Some authors even argue that it is not necessary to find a solution to such a debate, because the final answer would not change what the purpose of artificial intelligence should be. Not that of being able to reproduce human intelligence, but simply to facilitate the lives of men as the maximum expression in the field of automation.

[103]R. Searle, “Minds, Brains, and Programs,” Behavioral and Brain Sciences 3, no. 3 (1980): 417-424.

[104]Philip C. Jackson, Introduction to artificial intelligence, Dover Publications, 2019, XX.

A study conducted in 2016[105]it also revealed that a large percentage of artificial intelligence scholars are convinced that no research will ever be able to make a machine completely reproduce the mechanisms of human intelligence.

However, it is also worth mentioning a position, particularly complex and for which reference is made to the texts dedicated to it, which has developed in recent years among a small group of scientists.

In particular, this different approach is based on the analysis of a given system to verify whether the latter is capable - at least in principle - of supporting "human level intelligence".

The latter is described throughthe so-called high level mentalites – abilities such as language understanding, imagination, and consciousness – not yet reached by artificial intelligence systems.[106]

In summary, to draw a line on the issue, the opinion of the majority of the scientific community is shared according to which the AI systems developed to date are not yet capable of "thinking" like a human being.

To date, the previously analyzed machine/deep learning systems have led to formidable achievements - it is difficult, to be honest, not to be deeply impressed by GPT-3 - but they have not yet allowed the creation of an "intelligent" machine. ", rather only of good imitators/automatons.

In the opinion of the writer, without having the presumption of resolving a debate that is not yet dormant but in order to give a specific "cut" to this thesis, the intelligence of a system/machine should be reconnected to what is commonly defined "free will."

A machine can be said to be intelligent when it is capable of making autonomous decisions, which come from the teachings of its programmer and are not actually predictable by the latter, because they are the result of an independent decision-making capacity.

Which notimplies that they are able to think like or better than a human being, because even an animal, despite an "inferior" intelligence, is capable of making autonomous decisions. One could almost say that a machine is intelligent the moment it is able to

[105]Vincent C. Müller and Nick Bostrom, "Future Progress in Artificial Intelligence: A Survey of Expert Opinion," Fundamental Issues of Artificial Intelligence, 2016, pp. 555-572

[106]Philip C. Jackson, Toward human-level artificial intelligence: Representation and computation of meaning in natural language, Dover Publications, 2019, 4-5

"disobey."

Well, this could be a first point of contact with a regulation in law of the liability regime of artificial "intelligence" - in the sense previously specified - because, for example, the Italian legislator has shown himself sensitive to a liability intelligent - like an animal - and the damage it causes.[107]

It is true that a machine is not equipped with that "vital breath" that is commonly attributed to animals, but the idea that inspired the legislator to protect the position of those who have suffered damage from the latter probably concerns the ability to escape the master's "supervision" and make "decisions" - respond to instincts? – autonomous.

In conclusion, as fascinating as the philosophical debate that dominates the artificial intelligence scene is, we must always keep in mind that the results achieved to date by this type of technology have eliminated most of the doubts that were raised in the 1960s and following on actual capabilities of the machines.

This could mean that it would be premature to exclude a priori the creation of an "intelligent" system, taking into account the incessant and frenetic pace with which technological progress proceeds - just compare today's achievements with those of ten years ago - which, within in a short time, it could make what is currently considered impossible achievable.

Therefore, on the basis of this premise, it is necessary to carry out some reflections in law on a (future/possible) "intelligent" machine/system and the next chapter will deal with elaborating these considerations regarding the "attributability" deriving from a possible liability aquiliana of such an invention.

[107]Art. 2052 cc "The owner of an animal or whoever uses it for the time in which he has it in use, is responsible for the damage caused by the animal, whether it was in his custody, whether it was lost or escaped, unless that proves the fortuitous event."

CHAPTER II

Legal responsibility

Legal responsibility consists of the obligation and obligations that are connected to the performance of acts by individual citizens, bodies and public administrations. In essence, every legal entity is legally responsible.

The objective pursued in this phase is to analyze the different legal regimes of non-contractual liability proposed in relation to Artificial Intelligence technologies.

In this sense, when alluding toto this last legal institution, we intend to refer to the unjust damage caused to a "member" by a system - be it a driverless car, a robot/android/automaton, an electronic/digital system - regulated by artificial intelligence technology .

When the latter commits an "intentional or negligent act" that causes unjust damage, who is required to pay compensation for the damage?

As you will see, the responsibility of such an AI system inevitably involves a plurality of other subjects, such as the owner of the machine, the manufacturer of the machine and the programmer of the machine, whose positions have been explored in depth by recent initiatives of the the European Union regarding this matter.

The latter, specifically, realizing the fact that today's level of technological development is not sufficiently advanced to allow the creation of a completely "autonomous" and intelligent machine, push towards the responsibility of these latter figures, suggesting that the machine should be considered a mere product.

Here, however, we will describe some fascinating hypotheses - which could be defined de iure condendo - relating to the possibility of regulating the position of a future and possible machine that presents the aforementioned characteristics.

As we will see, the most "optimistic" legal scholars have attempted to equate such artificial intelligence technology with a real electronic person, configuring it as a legal entity, endowed with its rights and duties, with all the repercussions that this could have on any form of non-contractual liability.

Others, however, have proposed the analogical application of art. 2047 of the civil code

and art. 2048 cc, on the one hand, recognizing the "effective intelligence" of such a machine and on the other, noting the need to consider it at the same time as an "incapable subject" so as to ensure the involvement of further "responsible" subjects.

Still others, not believing that the threshold of legal subjectivity has been exceeded, suggest going back to the institutions of Roman law, comparing the position of such an AI system to that of the servus in ancient Rome.

In the same way, finally, always on the basis of the consideration of such a machine as an intelligent res, the possibility of applying the art analogically was put forward. 2052 cc regarding damage caused by animals.

For ease of exposition, we could take as an example a hypothetical - but perhaps not so unlikely - future intelligent robot, tasked with carrying out assistance activities in hospitals or providing support in centers for the elderly.

Before we can proceed with the in-depth analysis of these reconstructions, it is necessary to be clear about some fundamental concepts, which will be described briefly in the following paragraphs.

2.1 The concept of "subjectivity"

In order to fully understand the possibility of configuring a machine as an independent subject on the "legal scene" of relations between associates, it is first necessary to have a clear understanding of what the Italian legal system means by "subject of law" and how this can influence its possible form. of non-contractual liability.

To do this, it is necessary to start from article 1 of the Civil Code of 1942, the first paragraph of which states: "Juridical capacity is acquired from the moment of birth."

But what is legal capacity? The definition found in the main private law manuals is "the suitability to be the holder of subjective legal situations" and therefore to be considered a subject of law by Italian law.

Well, notthere is a summary rule of the legal entities that are recognized by the law, but it is very possible to identify them on the basis of a series of provisions contained within the code.

In fact, despite the art. 1 of the Civil Code seems to imply the recognition of such a capacity only for natural persons - understood as human beings - over time this has also been extended to various other figures.

First of all, legal capacity is also attributed to "entities", for which, however, it is necessary to make a traditional distinction: on the one hand, there are entities, such as joint-stock companies, which are considered "legal persons" because they have perfect patrimonial autonomy - that is to say that they are owners of their own assets which constitutes the generic guarantee pursuant to art. 2740 cc[108]–; on the other hand, entities such as unrecognized associations that lack a similar characteristic - the so-called "legal personality" - and, consequently, do not fall within the category of "legal persons."[109]

Although this is not clear, in recent years there has been a tendency to recognize a limited legal capacity even for such entities, especially when due to their organization they are particularly distinguishable from the natural persons who manage/train them.

This opening has been further extended by jurisprudence,[110] which identified the

[108]Art. 2740 Civil Code "The debtor is responsible for the fulfillment of the obligations with all his present and future assets."

[109]Andrea Torrente, Piero Schlesinger, Manual of Private Law. Twenty-second / edited by Franco Anelli, Carlo Granelli ed., Milan: Giuffrè Francis Lefebvre, 2015, 96.

[110]v. Cass. Civ., sec. a., sentence no. 19663/2014 which speaks of a "configurability within the condominium of an albeit attenuated legal personality, and in any case certainly, in progress, of an

possibility of attributing some subjective legal situations - active and passive - also towards "organised structures" (not for this reason entities) such as a condominium.

Secondly, in recent years the possibility of attributing legal capacity also to the figure of the unborn child - i.e. the "conceived" embryo awaiting the birth event - has made its way into doctrine and jurisprudence, sparking a profound and heated debate. not yet completely dormant.

In particular, two opposing positions have developed. On the one hand there are those who believe that the unborn child "is in any case endowed with autonomous legal subjectivity (specific, special, attenuated, provisional or partial, whatever you prefer) because he holds, on a substantial level, some direct personal interests, such as right to life, the right to health or psycho-physical integrity, the right to honor or reputation, the right to personal identity [...]"[111]

On the other hand, there are those who argue that the unborn child must be freed from the idea of legal subjectivity, and should rather be considered an "object" of the law, i.e. the something protected by the relevant regulations - think of the law on abortion , but also to the rights granted to the unborn child in matters of succession - without having to bother with the attribution of real legal capacity.[112]

Beyond the reasons that have pushed doctrine and jurisprudence to take one or the other position, the fact that the opportunity of conferring limited legal capacity on the unborn child was assessed demonstrates how the introduction of a new figure - which goes beyond the traditional two-way division of natural persons and legal persons - cannot in principle be excluded a priori.

In this sense, on the basis of a profound awareness on the topic that has occurred in recent years, the hypothesis of also considering animals as subjects of law has been put forward.

This solution is accepted at an international level - just think of other countries such as Switzerland, where the subjectivity of animals is recognized in a constitutional law,[113]and the Conventions onanimal rights - and embraces the idea that other living beings besides

autonomous legal subjectivity."

[111]Cass. Civ., sec. III, sentence no. 10741/2009.

[112]Tommaso Gazzolo, “The damage «before» birth”, Politics of law, 4, 2019, 614.

[113]Francesco Paolo Traisci and Fiore Fontanarosa, Animal rights: from objects of agri-food consumption to legal entities with their own rights, in Cibo e Diritto, ed. Lucia Scaffardi and Vincenzo Zeno Zencovich, RomaTRE Press, 2020, 863-864.

humans can also be entitled to some subjective legal situations.

The starting point of such a theory is the recognition, supported by the results of numerous research, of the ability to behave "rationally" and experience emotions within the animal world.

In fact, animals manifest in this way what is commonly defined as "vitality" and which represents the element in common with human beings.

Starting from this assumption, therefore, a part of the doctrine maintains that we should recognize at least a limited catalog of "basic" rights, such as "the right to life, the right to freedom, the right not to suffer"[114] for this category of subjects.

However, it must be highlighted that this reconstruction has remained a "minority" current of thought compared to the majority of authors, who share a more anthropocentric vision of law.

In fact, one of the main criticisms - to be honest, not particularly convincing - that is leveled at the attribution of the legal subjectivity of animals is of a strictly ontological nature: the attribution of a possible ownership of rights would be possible only if the animals were actually able to understand and perceive the norms, which are a product of human society for man (understood as a species).[115]

This observation is then joined by those who affirm the impossibility for an animal, even if it were recognized as having any rights, to have the necessary means to concretely assert them.

These are, in reality, criticisms devoid of any argumentative basis other than mere human discretion, to which one can easily respond that our system protects the situation of the incapable and minors, two figures who by definition may not be able to understand the rules fully and do not have the possibility to take action - independently - to protect their rights.[116]

Nonetheless, the jurisprudenceItalian would seem to settle on this last position, where the Supreme Court of Cassation has established that: "The animal, although it is a sentient being, cannot be the subject of rights for the simple reason that it lacks the so-called "legal capacity" (which is defined, precisely, as the ability to be the subject of rights and

[114]Diana Cerini, Law and animals: private law notes, Giappichelli Editore, 2012, 20.

[115]Gianpiero Paolo Cirillo, Institutional System of Common Law. Second ed., Wolters Kluwer, 2021, 230.

[116]Diana Cerini, Law and animals: private law notes, Giappichelli Editore, 2012, 20

obligations); capacity that the law reserves for natural and legal persons."[117]

Indeed, in the sentence in question, the judicial authority seems to translate that concept previously analyzed in reference to the figure of the unborn child, according to which an approach aimed at the objectification of this category should be adopted.

Specifically, animals should be considered "objects of law" or rather the recipients of the protection granted by the relevant regulations and not the holders of the rights themselves.[118]

This is an approach in line with the traditional "humanized" vision of law, although, in the opinion of the writer, it would seem to be constructed more to avoid the emergence of legal-philosophical questions regarding equalization or in any case rapprochement between beings humans and animals that on the basis of proven arguments.

For the purposes of this essay, however, it is important to underline how some of the same authors who are critical of the comparison between man and animal would seem to take a more lenient position towards a possible "autonomous digital agent" (center of attribution of rights and duties) on the basis of the fact that it would likely be able to reconstruct its own decision-making process, because, unlike animals, it would be "produced by man according to human intellectual patterns."[119]

The suitability to be the owner of subjective legal situations (ie legal capacity), however, does not automatically imply the recognition of the ability to independently look after one's own interests and, indeed, must be kept distinct from the ability to validly carry out acts having legal effect.[120]

This last faculty, in fact, to the sensesof the art. 2 of the Civil Code, is called capacity to act and, in our system, is acquired upon reaching the age of majority, set at the age of eighteen.

Well, without going into the details of this institution which over time has undergone various evolutions due to the succession of different criteria to establish when a person

[117]Cass. Civil section II, sentence no. 22728/2018.

[118]Ibid., the Court's reasoning concludes with: "the common expression "animal rights" must be understood in a non-technical, juridical sense, meaning it refers, not to the (inconfigurable) ownership of subjective rights by animals, but to the complex of legal protection that public law provides in defense of those living beings."

[119]Gianpiero Paolo Cirillo, Institutional System of Common Law. Second ed. (Assago: Wolters Kluwer, 2021, 231.

[120]Paolo Zatti, Vittorio Colussi, and Arianna Fusaro. Outlines of Private Law. Eighteenth ed. (Wolters Kluwer, 2020), 149.

was actually able to take care of himself, here it is interesting to highlight the presence in our system of figures who are considered legal subjects, but incapable of acting.

In particular, these are subjects - natural persons - who in the eyes of the law appear to be in need of special protection because they are considered individuals who can be easily attacked or deceived due to their inability to manage their interests adequately, as demonstrated by the heading of Title XII of the Civil Code which contains the discipline: "Measures for the protection of persons deprived in whole or in part of autonomy."

Every "person without autonomy" has a specific regime, the rigidity of which is parameterized on the basis of the reasons that lead the system to build a "protective cage" around them, and yet it is possible to trace the common thread that unites these subjects passing through two elements: (i) a support figure who supports the incapable person in carrying out acts with legal efficacy; (ii) the regime of annulment of the latter, ensuring the possibility for the incapable person to "retrace their steps" and thus avoid prejudice.[121]

Reasoning in a broad sense, therefore, the "incapable" could be seen as sentient/intelligent beings to whom it is unimaginable to deny the ownership of rights, but who, at the same time, the legislator does not consider capable of fully understanding the "weight" of their actions and who are therefore not granted such "negotiating capacity."

On the basis of this last statement, we begin to see the first similarities with the situation of a future "intelligent" robot and, for this reason, we need to ask ourselves whether and how the Italian law regulates the damage causally attributable to the conduct of an unable.

Before continuing, however, it is necessary to provide a brief reconstruction of non-contractual liability also in order to better contextualise the investigation plan that is intended to be carried out.

[121]Paolo Zatti, Vittorio Colussi, and Arianna Fusaro, Outlines of Private Law, Eighteenth ed. Wolters Kluwer, 2020, 150-151.

2.2 – Civil liability

As can be imagined, rivers of ink have been dedicated to non-contractual liability and the commitment of the most authoritative scholars from the past to today has meant that an endless literature has developed on the point.

A book would not be enough to highlight and explore the countless aspects highlighted by the detailed analysis of this institute and, therefore, here it is interesting to outline the essential characteristics of the latter that intersect with the theme of this thesis.

The institute in question, as is known, responds to the need to give a concrete transposition to the fundamental principle of neminem laedere, expressed differently depending on the historical eras.

In fact, the origins of the "Aquilian" responsibility date back to the era of Roman law and precisely to the approval - around 286 BC - of the lex Aquilia de damno with which the figure of the damnum iniura datum was introduced to the crimes already foreseen by the law of the XII tables.[122]

Specifically, the damnum iniura datum was intended to protectthe damage - usually referable to the killing of a slave or another's animal - caused dolo aut culpa, i.e. intentionally or following behavior attributable to incompetence, negligence or imprudence.

Precisely the subjective element of guilt has played a primary role in the development of the modern concept of civil tort, constituting the fundamental basis for the emergence of concrete responsibility.

As evidence of this, we can observe how the teachings of Roman law were first taken up by the French Code Civil, whose art. 1382 is expressed in the following terms: “Tout fait tantoconque de l'homme, qui cause à autrui un dommage, oblige celui par la faute duquel est arrivé à le réparer”.

In the same way, then, the Italian Civil Code of 1865, literally translating the provision of the Napoleonic code, identified in art. 1151 this rule: “Any act of man that causes damage to others obliges the person through whose fault it occurred to compensate for the damage.”

However, with the provisions of the Code Civil and the Civil Code of 1865 the transition

[122]Antonio Masi, Lessons on institutions of Roman law. CEDAM, Padua, 2012, 139.

began to be outlined from the conception of the typicality of the tort of Roman law - understood as a preliminary identification of the anti-legal conduct that would have entailed the obligation to compensate for the damage - to the development of a general civil liability clause "aimed at providing relief to a multiplicity of interests, to be identified from time to time in the interpretative process (so-called atypical nature of the offence)"[123] founded on the subjective element of the acting subject.

As numerous scholars in the sector have pointed out, the centrality of guilt was due to the conception of the tort as a transposition of the "crime" within the civil context.

Like the criminal offense, therefore, it was necessary to demonstrate that the person "responsible" for the civil offense was actually "guilty" or that the damage could be "reprimanded" on the basis of his conduct.[124]

As a consequence of such a comparison, the idea initially developed that the rules of non-contractual liability had a sanctioning function, aimed at repressing the conduct of the tortfeasor which damaged the legal sphere of others, taking advantage of the freedom granted to him by the 'sorting.

Even the main rule on the subject of Aquilian liability of the Civil Code of 1942 is the result of this approach, as evidenced by the provisions of the art. 2043 cc, pursuant to which "Any intentional or negligent act which causes unjust damage to others obliges the person who committed the act to compensate the damage."

Asin fact we read in the Report of the Minister Keeper of the Great Seals to the Civil Code of 1942: "for the intentional or negligent act to be a source of liability it must produce unjust damage. It is thus clarified [...] that culpa and iniuria are distinct concepts; and therefore it is required that the fact or omission, to be a source of liability, must be intentional or negligent, i.e. attributable."

Yet, despite the indispensable character enjoyed by the subjective element of guilt, the legislator of 1942 had already prepared a series of alternative cases

- which, with technological progress, have increasingly taken on greater importance - in which this criterion did not play such a central role but, on the contrary, appeared almost left aside.

These are hypotheses that have been recalled by legal experts in the context of the

[123]Paolo Cendon, Civil Liability. UTET juridical, Turin, 2020, 25.
[124]Cesare Massimo Bianca, Civil law 5. Responsibility, Giuffrè Francis Lefebvre, 2021, 522-523.

analysis of the non-contractual liability of an artificial intelligence system and, for this reason, it is necessary to provide an initial description of the rationale behind them.

The advent of "industrial society" has favored the proliferation of situations likely to cause damage - think for example of the circulation of an ever-increasing number of cars - in reference to which the attribution of negligence did not seem sufficient to guarantee compensation of the injured party.[125]

The legislator of 1942 seems to have regulated the aforementioned innovative situations in a different way compared to the general clause of the art. 2043 cc, realizing the intrinsic danger: i) of some activities, defined as dangerous by article 2050 cc which regulates the liability regime, but also of the "Ruin of a building" pursuant to art. 2053 and of the "Circulation of vehicles" pursuant to art. 2054 of the civil code; ii) the use of certain "things", dealing with "Damage caused by things in custody", as stated in the heading of the art. 2051 cc; iii) the instrumental use of animal intelligence, governed by the aforementioned art. 2052 cc; iv) the instrumental use of the human intelligence of others and in particular the assignment of domestic servants and clerks by masters and clients pursuant to art. 2049.[126]

To this last predictionthe provisions of the art. can be combined. 2047 cc and 2048 cc, which respectively regulate the "damage caused by the incapable person" and the responsibility of parents, guardians, tutors and art teachers", since these are rules that involve a person other than the person has materially committed the act and therefore they are traditionally called "vicarious liability" or for the actions of others.

The peculiarity of all the articles mentioned above - with the exception of the articles. 2049, 2053 and part of the art. 2054 cc - consists in the provision of a possible exculpatory proof in favor of the injurer, or rather the person called to respond, on whom however the honor of proof rests.

In other words, where the latter person is not able to demonstrate in court the liberating element envisaged by the provision applicable to his conduct, he will be sentenced to compensation for the damages caused.

Now, in reference to the presence of the liberating proof, two different orientations have

[125]Francesco Galgano, Treatise on civil law, CEDAM, 2014, 129.

[126]Ugo Ruffolo, Artificial Intelligence and Responsibility: Responsibility From Algorithm? ; AI and Self-Driving Cars, Production Automation, Medical-Pharmaceutical Robotization; AI and Contractual Activities; Union Trends and Disciplines, Giuffrè Francis Lefebvre, 2017, 13

developed:

i) The first orientation attempts to remain anchored to the fundamental principle of the necessary presence of the subjective element, equating the provision of a liberating test with an inversion of the onus probandi of guilt, as if in such hypotheses guilt was presumed. The exculpatory proof, in this sense, could theoretically be limited to the demonstration of the absence of fault and, in other words, the injurer should demonstrate that he has not committed the sin of incompetence, negligence or imprudence.

Under the lens of this approach, the articles previously examined - again excluding art. 2049, 2052 and part of 2054 of the Civil Code which are considered the only hypotheses of strict liability, which will be discussed shortly - are interpreted as hypotheses of aggravated liability, where the burden consists precisely in the reversal of the burden of proof.

ii) The second orientation, which is becoming increasingly popular in recent years, interprets the various liberating tests not so much as a demonstration of the absence of guilt, but rather as an exemption foreign to the sphere of the subjective element, which requires something more than to the simple absence of guilt. The articles in question are therefore described as hypotheses of objective liability, because they ignore the presence of fault and the subject is called to respond for the simple fact that his conduct can be subsumed in the foreseen case.[127]

Without entering into the lively debate on the applicability of each single rule mentioned above within the category of aggravated liability or that of objective liability, what is important to underline is that both orientations start from the precise aim of favoring, in a different way, the possibility that the injured party can obtain compensation.

It is clear that a greater objectification of responsibility and the consequent more rigid regime ensure greater protection for the injured party, worsening the position of the injurer (or that of the person responsible for the actions of others) who will have a greater chance of being held accountable of the damage caused.

However, both approaches demonstrate the majority doctrine's move away from the idea according to which the rules relating to non-contractual liability have a mere

[127]Cass. Civ., sec. III, sentence no. 15383/2006 (subsequently taken up by Cass. Civ., section III, sentence no. 4279/2008: "in strict liability the judgment is purely typological and consists in ascertaining whether or not the event that occurred belongs to the series of those that the criterion of attribution ascribes to a certain sphere of the subject due to their simple occurrence. In these terms the centrality of the causal link in the hypotheses of objective responsibility is exact."

sanctioning/deterrent function, in favor of a conception focused on the compensatory nature of the same, which have as their objective the last one is to keep the injured party harmless from the losses suffered.[128]

On the basis of these reflections and the consequent wavering of the criterion of fault, part of the doctrine has developed various theories trying to reorganize the system of rules regarding non-contractual liability, which find their common foundation in the concept of "risk" as a criterion of attribution .[129]

A first thesis[130]describes objective liability as the liability deriving from "business risk", i.e. from the carrying out of socially useful activities, which therefore could not be prohibited a priori, but which at the same time represent a risk for the community due to their intrinsic danger.

The fundamental core of such an approach is found in the idea that the construction of a more rigid system of responsibility that does not take into consideration the subjective element pushes the subject to whom the organization of such activities refers to adopt the most suitable to avoid the occurrence of damage.

Furthermore, it is worth reiterating for the purposes of this thesis, it entails a greater guarantee for the injured party of not being forced to bear the damage suffered, making the position of the injurer more onerous.

The Author uses the term "business risk" because most of the aforementioned forecasts refer to the carrying out of entrepreneurial activities, but, as described previously, such reasoning also extends to activities which, in common opinion, involve a high risk for other associated companies (for example, the circulation of vehicles).[131]

In this vision, therefore, strict liability goes hand in hand with fault liability, having to refer to this concept only in specific cases and leaving the general clause of art. 2043 of the Civil Code intact.

Yet, this orientation has not been free from criticism: some authors contest the use of an attribution criterion based mainly on economic and non-legal reasoning, while others focus on the disparity in treatment that could be created between small and medium-

[128]Mario Barcellona, Treatise on Civil Liability (UTET juridical, 2011), 43-44. But see also Guido Alpa, Civil liability, (UTET juridical, 2018) where reference is made to the concept of "compensation", pages. 46 et seq.
[129]Ferrari V., New profiles of insurance law. The insurance fact, Giuffrè, 2003.
[130]Pietro Trimarchi, Risk and Objective Responsibility, Giuffrè Editore, 1961.
[131]Paolo Franceschetti, Civil liability, Maggioli Editore, 2009, 30-31.

sized businesses, which would be most affected by the rigid liability system, and large multinationals.

A second thesis[132]instead insists on the loss of centrality of the criterion of guilt, proposing the criterion of created risk as a substitute.[133]

In summary, if an individual carries out an activity - not necessarily dangerous or "business" - capable of giving rise to a risk of causing harm to others, then he must be held accountable for any prejudice.

This approach particularly enhances the ancient Latin phrase “cuius commoda eius incommoda”, often referred to in the criminal law field, for the explanation of which we refer to the words used by the Supreme Court of Cassation: “he, in whose favor an action is carried out activity, bears the risks inherent in carrying it out and therefore also suffers the effects of any harmful consequences."[134]

The underlying objective isin reality twofold: on the one hand to make the civil liability rules perform a deterrence (prevention) function, on the other to ensure that the ones most suitable to bear the loss are responsible for the damages according to the needs of justice and the economic structure of interests.[135]

The peculiarity of this theory lies in the fact that the criterion for attributing the risk created actually modifies the very concept of guilt and is therefore applicable, unlike the school of thought previously described, also to cases attributable to art. 2043 cc

Specifically: “the fault is no longer*the negative qualification of the fact, but only the relationship between the severity and probability of the harmful event and the cost necessary to avoid it."*[136]

On the basis of this single criterion for responsibilityin extra-contractual terms, in conclusion, anyone who would have the possibility to prevent the damage - for example by not carrying out his activity or behaving differently - and who does not do so, is required to answer for the damage caused.

Finally, for completeness of explanation, it is necessary to point out a third position on the subject of extra-contractual liability, attributable to the vision of a part of Italian

[132]Guido Calabresi, Cost of accidents and civil liability, Giuffrè, 1970.
[133]Paolo Franceschetti, Civil liability, Maggioli Editore, 2009, 27.
[134]Cass. Pen., sec. IV, sentence no. 38154/2009.
[135]Carlo Castronovo, The New Civil Responsibility, Milan: Giuffrè, 2006, 280.
[136]Paolo Franceschetti, Civil liability, Maggioli Editore, 2009, 28.

doctrine which does not agree with the identification of a single criterion of attribution for civil torts, be it negligence or it is the risk created.

There would therefore exist a plurality of criteria for attributing responsibility to the Italian legal system, starting from guilt for the hypotheses regulated by the art. 2043 cc - for which an almost residual position emerges - up to the specific provisions of the code and extra code, the regulation of which is modeled on the attempt to ensure greater guarantees for the injured party by using a different attribution criterion.

Now, as will be seen better in the next chapter, the development of a regime of extra-contractual liability of an artificial intelligence system cannot ignore the search for a balance between the position of the subject possibly called to respond for damage to a machine and the position of the injured party whose losses must be repaired, and therefore between the lesser and greater "objectification" of responsibility.

AOnce this introductory excursus has been completed, it remains to contextualise the general clause of non-contractual liability by adding the further element of imputability to the equation, before reflecting on the possible hypotheses of legal regulation of the damage caused by an "actually intelligent" machine.

Doctrine and jurisprudence agree in qualifying the provision of art. 2043 of the civil code as a case with a complex structure, although there is still some timid discussion regarding its constituent elements.[137]

Traditionally, however, the majority opinion has always supported the following structure, also endorsed by the Supreme Court of Cassation: “The constituent elements of the Aquilian offense are the conduct, the psychological element, the unjust damage and the causal link. It follows that, if the judge deems any of these elements to be non-existent, the request for compensation for damages must be rejected without the need to ascertain the existence of the others".[138]

Without going into the merits of the description of each individual constituent element, what is interesting to underline here is precisely the reading of the art. 2043 cc in light of the provisions of the art. 2046 cc entitled "Imputability of the harmful fact".

[137]Guido Alpa in Civil liability, UTET juridical, 2018, 24 et seq., underlines how the debate concerns the number of requirements or constitutive elements, two schools of thought being able to be identified: "according to Stolfi there are only three: (i) violation of a right of another; (ii) subjective responsibility (fault or willful misconduct); (iii) harm. According to De Ruggiero, there are four: (i) violation of the law; (ii) positive behavior; (iii) damage; (iv) voluntariness and imputability.”

[138]Cass. Civ., sec. III, sentence no. 2422/2014.

This article, which constitutes a step forward compared to the 1865 Code, provides that the agent must be able to understand and want - at the moment in which the act is committed - in order to be held accountable for the damages of his action .

Although for a long time the doctrine has considered the notion of imputability as perfectly overlapping with that of guilt, as evidenced by the passage of the Report to the above-mentioned Code, however for many years now the majority opinion has been in favor of a distinction between them.

In particular, the concept of fault should be traced back to a "deviation from a behavioral standard that also considers the concrete situation in which the agent finds himself" and is in all respects a constitutive element of the case.

The notion of imputability, however, rather concerns the possibility that the agent is affected by a psychic disturbance, which exempts him from being held responsible.[139]

If we analyze art. 2046 of the Civil Code, reasoning a contraris, we can easily deduce that the ability to be held accountable for the damages caused by one's conduct (so-called criminal or non-contractual capacity) does not depend on whether or not the agent has the capacity to act, but rather only by the fact that he has the capacity to understand or want at the moment of carrying out the act.

Consequently, it is very possible that one of those subjects deemed "incapable" by the law may remain obliged to pay compensation for the damages resulting from their action, provided they meet the required requirements.

Specifically, the jurisprudence of the Supreme Court of Cassation requires that the judge "must ascertain on a case-by-case basis whether, in relation to age, psycho-physical development, the manner of the crime or other reasons, the capacity to to understand or want."[140]

In practice, however, a subject who the legal system deems incompetent, even if endowed with the ability to understand and will - to be understood as the ability to realize his own conduct and self-determine - cannot be called directly to answer and, for this reason , the legislator involves additional subjects who should supervise the incapable person.

About that, in fact the legislator has prepared the aforementioned art. 2047 cc, which talks

[139]Pasquale Stanzione, Non-contractual liability. Volume Two, CEDAM, 2012, 64 et seq.
[140]Cass. Civ., sec. III, sentence n 11163/1990.

about the figure of the supervisor and art. 2048 cc, which instead refers to parents, tutors and art teachers, subjects who should be responsible for the education, in a broad sense, of their children, apprentices and students.

At this point, once the fundamental concepts of the institutions of legal subjectivity and non-contractual liability have been outlined, we finally have the necessary tools to deepen the analysis of the different theories proposed in relation to a possible future intelligent machine.

2.3 – Towards an "Electronic Personality"

As anticipated in the first paragraph of this chapter, some legal experts have suggested the development of a new category of legal entities that complements the already known categories of natural person and legal person, namely the "electronic person."

The possibility that in the future humanity will have the necessary technology for the creation of a machine so intelligent that it can, if not be equated, then at least brought closer to a human being, is also implicitly recognized by the "insiders" of the European Union.

So much so that, as stated in point 59 of the European Parliament Resolution of 16 February 2017 - Rules on the civil law of robotics - the Commission was invited to consider and evaluate "the establishment of a specific legal status for robots in long term, so that at least the most sophisticated autonomous robots can be considered as electronic persons liable to compensate for any damage caused by them, as well as possibly the recognition of the electronic personality of robots that make autonomous decisions or that interact independently with third parties."

But what would be the consequences of attributing legal capacity to a machine - the intelligent robot that assists the elderly in our example - and, consequently, of considering it the owner of subjective legal situations?

First of all, we should be willing to recognize the existence of rights and duties on the part of the artificial intelligence system.

As dystopian as it may seem, the equation between a physical person and an electronic person would entail the subjection of the latter to the principles established by the Constitution and the fundamental rights contained therein.

It is a statement certainly destined to cause discussion, because if on the one hand it is clear that the robot in our example could boast the right of ownership over certain assets and, consequently, be the owner of an asset that would constitute its generic guarantee pursuant to art. 2740 cc, on the other hand it is more difficult to agree on the possibility for the machine to boast typically "human" rights.

Reference is made to the right of personal freedom, to the right of free movement (in general, therefore, the Civil Relations of Title I of the Constitution) but also to political rights, which offer ideas for a debate bordering on science fiction but which is possible

to start arguing.[141]

The legal subjectivity of the machine governed by artificial intelligence technology would not only imply the attribution of rights, but also the assumption of specific duties, such as for example those deriving from the principles of solidarity and contribution to common well-being.[142]

In this sense, can the robot in our example, after being paid for the activity carried out, be asked to pay taxes?[143]

It is clear that the development of the concept of the electronic person poses a whole series of questions on which it is possible to speculate at length, but which at the same time dictate the outline of the major questions that deserve to be addressed in the coming years.

However, this is not the caseplace to delve deeper into these fascinating topics, having rather to focus the discussion on the possible repercussions in terms of non-contractual liability deriving from the possible attribution of legal capacity to a machine.

In this regard, electronic persons, where they are recognized as holders of rights and duties, could be called to respond directly for the damage caused by their conduct in the same way as a natural person.

The robot in our example, therefore, would be subject to the previously analyzed disciplines of strict liability - where compatible with its situation - or, in general, to that of art. 2043 cc

Sticking to this last hypothesis, taking the former imputability for granted. art 2046 cc in the sense previously described and therefore that the machine is capable of "understanding and wanting", the first difficulty arises precisely in relation to the subjective element of fault.

How is it possible to demonstrate the imprudence, incompetence or negligence of an intelligent machine, if often we are not even able to understand how it arrived at making a certain decision?

A possible solution to this problem could be offered by the implementation of that

[141]Alfonso Celotto, Can robots have rights?, in BioLaw Journal. BioDiritto Journal, 1, 2019, 92.
[142]Above all, the art. 53 Constitution, first paragraph: “Everyone is required to contribute to public expenses based on their ability to pay.”
[143]Jakob Wegener Friis, Should Robots Pay Taxes?, in Shaping Europe's digital future - European Commission, February 27, 2019.

ethical principle of transparency indicated in point 12 of the aforementioned Resolution of the European Parliament of 16 February 2017 and in particular the provision for all robots of "a "black box" that records the data about every operation performed by the machine, including the logical steps that contributed to its decisions."

The transparency desired thereinreality represents a point of contact with the discussed right to explanation - hypothesized regarding the protection of personal data with EU regulation 2016/679, better known as GDPR, to which the resolution in question also refers - "i.e., the fact that a those who are impacted by decisions of artificial intelligences are owed explanations as to how these artificial intelligences arrived at these decisions."[144]

Beyond the arguments for or against the existence of such a right, what is of interest here is the food for thought offered by Busto regarding the conclusions reached by researchers at the Alan Turing Institute.[145]

In particular, the aforementioned scholars suggest the creation of a "trusted third party" that is able to offer interested parties an adequate explanation of how the AI systems that process their data have reached the conclusion that it has had repercussions on their legal sphere.

Translating this approach into the context of the direct non-contractual responsibility of an artificial intelligence system, one could almost think of entrusting it to a future (and for the moment also difficult to imagine) trusted third party or in any case to a specific commission/committee for the reconstruction of fault within the illicit conduct of a machine.

This is clearly a hypothesis of fantasy law and, regardless of the problems regarding the formation and "jurisdiction" of such a body, in the analysis of such an eventuality one cannot fail to take into consideration the difficulty on the one hand in the implementation of the aforementioned " black boxes" and on the other in the reconstruction of a machine's decision-making process, not only for ordinary people, but also for those who taught it how and what to learn.

Precisely this last argument testifies to the tortuosity of the use of the already existing

[144]Nicola Busto. The electronic personality of robots: risk management logics between transparency and trust, in Ciberspazio e Diritto, 59, 3, 2017, 505.
[145]Sandra Wachter, Brent Mittelstadt and Luciano Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation," International Data Privacy Law, 7 (2, May 2017): 76–99

categories of non-contractual liability, parameterized mainly for natural persons.

Which leads us to suggest the development, together with the attribution of legal capacity to an AI system, of a specific regime for civil torts, which takes into account the inevitable differences between man and machine, inevitably characterized by an objectivisation of responsibility, for the sole fact of "being a machine."

Such a regulatory intervention, moreover, would have an epochal significance and, as such, should hopefully be launched at a global level, in order to avoid unequal treatment.

Already from these first lines, it is possible to realize how the conception of electronic personality and the consequent hypothesis of a "software agent"[146] was developed primarily for practical reasons.

It represents in factthe only liability regime that ignores the involvement of additional parties (owner, manufacturer, programmer of the machine), avoiding the risk of a "coverage gap" for the injured party resulting from the analogical application of the already existing categories of liability.

Nonetheless, the subjectivity of the machine would resultuseful for preparing a solution in the face of particularly complex cases, in which "active" proof of guilt on the part of the injured party or "liberating" proof on the part of any further responsible party appears difficult to demonstrate.[147]

However, if on the one hand such a solution would combine with the compensatory nature of non-contractual liability previously described, on the other hand its implementation would seem to cause more problems than it actually solves.

Just think of the aforementioned problem of the blameworthiness of the conduct of a robot - and the related problem of the burden of proof on the injured party according to the provisions of the art. 2043 cc - and that of the formation of a separate and autonomous heritage.

The main criticism leveled at the electronic personality, however, derives from a logical reasoning: although a future software agent is abstractly entitled to rights and duties deriving from legal norms, in reality the actual recipients of these precepts would still be

[146]Ugo Ruffolo, The problem of electronic personality, in Journal of Ethics and Legal Technologies 2, no. 1 (April 2020), 81.
[147]Erica Palmerini, Subjectivity and artificial agents: a solution in search of a problem?, in Observatory of civil and commercial law, 9 (2, 2020), 464.

the producers and the programmers of the machine itself, who should "educate" it about and respecting them.

In the event of non-compliance and consequent causing of damage, many would be willing to consider the figures mentioned above as the ones actually responsible, causing the attribution of legal capacity to the AI system to be short-circuited, which would therefore be meaningless .[148]

The European Economic and Social Committee also takes this position, according to which the introduction of a form of electronic personality "would entail an unacceptable risk of moral hazard"[149]and that only a natural person can be held responsible for damage caused by a machine.

Otherwise, the risk could be to "remove responsibility" from the natural persons behind the design of the AI system, thus eliminating the preventive function of correcting behavior (deterrence) typical of liability law.

The Opinion of the European Committee,as we will see in the next chapter, it reflects the position taken by the EU towards artificial intelligence systems, which are not seen as "sentient beings" but as real "products."

Of interest is the position regarding the vexed question of electronic personality.

Already at the beginning of the nineties[150]the possibility for hyper-intelligent systems to be qualified as subjects of law was reflected on, resolving the question in a negative way given the level of technological development of the time.

However, progress in artificial intelligence has expanded significantly in recent years, not least with self-learning and deep learning techniques.

Inevitably, over time, this exponential development has led to a necessary reconsideration of the question in question and the consequent philosophical problem of the transition from thing to person, also in light of the maneuvers and reflections carried out on the topic at a European level.

[148]Andrea Amidei, "Intelligent Robotics and Responsibility: profiles and evolutionary perspectives of the European regulatory framework" in "Artificial Intelligence and Responsibility: Responsibility from Algorithms? ; AI and Self-Driving Cars, Production Automation, Medical-Pharmaceutical Robotization; AI and Contractual Activities; Union Trends and Disciplines.", ed. Ugo Ruffolo, Giuffrè Francis Lefebvre, 2017, 98.

[149]Opinion of the European Economic Committee of 31 August 2017.

[150]Giancarlo Taddei Elmi, The rights of artificial intelligence between subjectivity and value. Ius condendum or fantadiritto?, in Il meritorious di tutela, Luigi Lombardi Vallauri ed., Giuffrè Editore, 1990.

According to Taddei Elmi, "the recognition of electronic personality" which is discussed in the aforementioned 2017 European Parliament Resolution, has been misunderstood by those authors who have advanced the equation between human subjectivity and that of the machine.

The Resolution, in fact, would not have proposed the recognition of an ontological subjectivity like that of natural persons, but rather would have proposed the attribution of an "ascriptive" subjectivity, in the wake of that attributed by our legal system for practical functional reasons to legal persons .

The Author insists on the dichotomy between "recognition" and "attribution" because it constitutes the true distinction between these two forms of subjectivity, which have an absolute and relative character respectively.

Consequently, from the point of view of liability, a machine could never be called to respond directly for the damage caused, except as an ascriptive legal person.

This position therefore seems to share the assumptions of a part of the German private doctrine, which finds its greatest exponent in Teubner,[151]according to which artificial intelligence systems should be described as flows of information so complex that they can be attributed, under certain conditions, a sort of "social identity."

Such "social identity", also to be understood as the possibility of being seen as an autonomous interlocutor by associates who come into contact with it, could constitute the reason for attributing partial legal subjectivity to the machine.

The latter, however, would have a fairly slight impact on the non-contractual liability of the AI system, since it would not exclude the involvement of the individual behind the software agent.

As Taddei Elmi points out, in fact, Teubner supports a particular variant of the liability of auxiliaries, which in our legal system is regulated by art. 2049 cc

Nonetheless, the Author does not close the door to the possibility that in the future a machine could develop the characteristics of ontological subjectivity - in particular consciousness - and at that point, it would become very difficult not to attribute the qualification of subject to this intelligent system as well.[152]

[151]Gunther Teubner, Digital legal entities? On the private status of autonomous software agents, translation by Femia P., Edizioni Scientifiche Italiane, 2019.

[152]Giancarlo Taddei Elmi and Francesco Romano, “The robot between ius condendum and ius conditum”,

This conclusion is based on some research by robotics experts, according to which consciousness is nothing other than a very high calculation capacity with which the human brain is equipped, so the reproduction of this within a machine would provide it of a conscience.[153]

Based on a statistical analysisof the time taken to develop increasingly complex memories, a machine could boast a computational capacity similar to that of humans as early as 2029.

In conclusion, the issue of "electronic personality", if not addressed in advance of the moment in which an AI system exceeds the threshold of subjectivity, could find systems around the world unprepared, with economic consequences (slowing down progress?) and above all legal.

Whether one wants to recognize an ontological personality or whether one wants to attribute an ascriptive personality, the fundamental problem to be resolved remains the attribution of an asset with which the machine itself or the subjects behind the new legal person can guarantee fulfillment of its obligations pursuant to art. 2740 cc

At the moment, however, such a problem appears insurmountable, since the development of a discipline in this sense would involve an epochal effort if contextualized at a global level, and, therefore, also the European Union, as we have already had the opportunity to explain, it is moving in a different direction.

Informatics and law, XXV, 2, 2016, 137.

[153]Giorgio Buttazzo, Artificial consciousness: mission impossible, Digital World 1.1 (2002), 16-25.

2.4 - Artificial Intelligence and responsibility for the actions of others

Art. 2047 cc

From the exposition contained in the paragraphsprecedents, it has emerged that our legal system also regulates the liability profiles for those legal entities - incapable of acting - in relation to the damage caused by their conduct.

In the event that one is not willing to consider a future intelligent machine sufficiently developed to be granted complete autonomy, but at the same time does not want to give up the recognition of its "sentient being", the solution could consist in the comparison between the subjects mentioned above and the artificial intelligence systems in question.

We should therefore imagine a machine that is "intelligent" but which at the same time cannot be left to itself, equating it with those subjects who, as specified in paragraph 2.2, the law deems necessary to protect through the involvement of figures close to them .

Pursuant to art. 2047 cc - entitled "Damage caused by the incapable person" - where a third party suffers damage caused by the conduct of a non-attributable person pursuant to art. 2046 cc, the damage must be repaired by those who are required to supervise the incapable of understanding and will.

The supervisor, however, has the possibility of freeing himself from responsibility by means of exculpatory proof, consisting in the demonstration of the impossibility of preventing the fact.

In this case, pursuant to the second paragraph of the article in question, the judge will be able to evaluate whether or not to condemn the incapacitated person to pay a fair compensation, taking into consideration the economic conditions of the parties involved.

It should be noted that the art. 2047 of the Civil Code constitutes, according to the majority opinion, a hypothesis of aggravated liability, where the guilt of the supervisor is presumed (i.e. reversal of the burden of proof) and the latter has the possibility of overcoming this last presumption by means of exculpatory evidence.

Traditionally, the case described in the article in question has been interpreted as liability for the actions of others, the supervisor not being the material author of the action, but, to

be honest, part of doctrine and jurisprudence.[154] recognizes you a direct responsibility deriving from the culpa in vigilando of the supervisor.

The duty of supervision of the latter derives from a qualified source, consisting either in the law (for example for parents towards minors) or from a contractual source, even if jurisprudence is slowly moving towards a more open position, starting to admit responsibility for “de facto surveillance”.

As regards the liberating test, however, over time a plurality of meanings have been attributed to it. A now-dating ruling from the Court of Cassation, however, seems to be satisfied with proof consisting in "demonstrating that this fact would have occurred in any case even if surveillance had been exercised, and therefore that there is no causal link between the omission of it and the harmful fact."[155]

Finally, having regard to the fair compensation provided for in the second paragraph of the art. 2047 cc, this is a provision in contrast with the principle expressed by the immediately preceding article, which establishes that the incapable of understanding and will is not responsible for the harmful consequences of his conduct.

However, the determination by the judge of any compensation calculated on a discretionary basis on the basis of equitable criteria responds to the need for "compensation" for the damage suffered by the third party mentioned above, because the latter can only be left without compensation in exceptional cases.

Now, trying to apply this system analogically to the situation of a "sentient" machine but incapable of understanding and wanting, if the latter engages in conduct that is harmful to the legal sphere of a third party, it is necessary to verify who could be the person required to monitor the car and called to compensate for the damage.

On the basis of what has been stated previously, the need for a specific source (legal, contractual, factual) from which the duty of supervision arises would seem to call into question the "owner" of the intelligent machine or in any case the subject who mainly benefits from the activity of the latter.

[154]Cass. Civ., sec. III, sentence no. 12965/2005: “The civil liability of the person required to supervise an incapacitated person, who has caused damage to third parties, derives from art. 2047 cc, which gives rise to a direct and proper responsibility of those who are required to monitor, for failure to comply with the obligation of custody, placing a presumption of responsibility against them, which can only be overcome by proof of not having been able to prevent the fact despite the diligent exercise of surveillance employed.”

[155]Cass. Civ., sec. III, sentence no. 5485/1997.

These subjects would become the supervisors of the intelligent machine when, through a contractual title, they assume control over it. The producer/seller/manufacturer of the machine, however, since the latter leaves his sphere of supervision when it is "entrusted" to another party, would remain extraneous to the events connected to the damage caused by it.

Of course, the idea that a sentient being can become someone's "property" evokes the concepts of slavery, which would inevitably conflict with the previously exposed constitutional guarantees that the artificial intelligence systems described should meet.

You should live with the idea that when you "purchase" the machine, you acquire a sort of "tutorship" and therefore have to monitor its correct behavior and prevent it from causing damage.

At this point, however, it is necessary to extend the reasoning to the disclaiming evidence of the guardian/owner of the machine, who would be burdened with the difficult demonstration of the absence of causal link between the lack of surveillance and the damage caused by the machine.

Which would translate into a totally favorable regime for the injured party, who would have excellent if not certain possibilities of retaliating against the person in charge of surveillance.

But even if the owner/guardian of the system governed by AI were somehow able to provide such liberating evidence in a trial, the judge would still see his equitable power provided for by the second paragraph of article 2047 of the civil code ablated, re-proposing the question of the possibility of attribute its own assets to the machine, which is the material cause of the damage.

As a result, the analog application of the 2047 cc does not seem to be entirely convincing.

Art. 2048 cc

The art. 2048 cc is strictly connected and connected with the art. 2047 cc where it identifies some typical figures of supervisors: parents/guardians, preceptors (ie teachers) and art masters, who are called to answer for the damage caused by the illicit act respectively of the minor children who live with them, of the students and of the

apprentices (also minors,[156]for the time they are under their supervision).

Yes, in this case, the parties held responsible are offered exculpatory evidence - it is anticipated, interpreted differently by jurisprudence with respect to the art. 2047 cc – consisting in the demonstration of not having been able to prevent the event.

A distinctive element between the two cases just described, which are therefore alternative to each other, concerns the capacity of understanding and will of the material author of the damage.[157]

Indeed, if the damage is caused by an incapable person (e.g. a minor), who due to a whole series of factors listed above (development, living conditions, age, etc.), does not have the capacity to understand and will, the 'art. 2047 cc, with the consequences described above. Otherwise, however, the case will be governed entirely by the art. 2048 cc

The art. 2048 cc is therefore considered a hypothesis of aggravated liability by the majority of legal scholars and, as such, is based on the presumption of guilt on the part of the recalled subjects, who are consequently burdened with the onus probandi.

However, this is a fault that differs depending on whether reference is made to the parents or to the tutors and art teachers: for both categories they will be liable for the failure to supervise the minor (ie culpa in vigilando), but with due regard for the parents jurisprudence has developed the concept of culpa in educando,[158] i.e. not having imparted sufficient education to the minor perpetrator of the offense and consequently having failed to fulfill the duties deriving from the art. 147 cc

All this has also had profound repercussions in terms of exonerating evidence for parents - going beyond the letter of the law - who, to overcome the presumption of guilt, will have to offer: "not the legislatively predetermined proof of not having been able to prevent the act (given that this is negative proof), but the positive proof of having given the child a good education and having exercised adequate supervision over him; all in

[156]Cass. Civ., sec. III, sentence no. 7387/2001: “The presumption of guilt referred to in art. 2048, paragraph 2, of the Civil Code cannot be considered applicable [...] in the case in which the student is an adult, it being presumed that, within the same provision, the legislator did not want to reserve for tutors and teachers art a worse treatment compared to that of the parents referred to in paragraph 1, extending their responsibility beyond the time limit of the minor age of the injurer."

[157]Manuela Mantovani, “The responsibility of parents, guardians, tutors and art teachers” in Commentary on the Civil Code of illicit facts articles. 2044-2059, ed. Ugo Carnevali, UTET juridical, 2011, 91.

[158]From this concept also arises the debated question - like the art. 2047 cc - if you are faced with a hypothesis of liability for the actions of others or with a hypothesis of liability for your own actions.

accordance with the social and family conditions, age, character and nature of the minor."[159]

The other figures contemplated by the rule in question, however, will have to provide as disclaiming proof the punctual fulfillment of their surveillance duties or the absence of a causal link between their lack of surveillance and the harmful event, also taking into consideration the predictability of this 'last.[160]

Well, from this first analysis of the art. 2048 it is well possible to realize the difficulty in applying the provision in question analogically without any type of adjustment.

In this sense, the responsibility of parents, with all the peculiarities connected to it, does not seem to adhere to the hypothesis of the responsibility of a system governed by AI technology, both because the "culpa in educando" cannot find a comparison in the possible owner/person who benefits from the activity of his business - as he is not the person who "educates" the machine - and because difficulties would arise in complying with the coexistence requirement.

The doctrine was rather focused onprovision of the second paragraph of the art. 2048, trying to highlight the position of the preceptor in his "abstract" conception of a figure called to instruct students.

In this way, it would be possible to bring out the responsibility of the figure of the programmer or in any case of the person who teaches the machine how to behave in certain ways - through the techniques described in the first Chapter - in relation to the damage that the latter causes.

It is clear that even in this case the case would need a correction to the limits of the de iure condendo.

In fact, it would be necessary to circumvent the provisions of the art. 2048 cc in which it is foreseen that the tutors are liable for damages caused by the students only during the time in which they are subjected to their supervision.

The machine programmer, once he has carried out his task as "educator", is deprived of

[159]Cass. Civ., sec. III, sentence no. 24475/2014. The Court's reasoning then continues: The inadequacy of the education imparted and the supervision exercised on a minor, the basis of the parents' responsibility for the illicit act committed by the aforementioned committed person, can be deduced, in the absence of evidence to the contrary, from the manner of the same act illicit act, which may well reveal the level of maturity and education of the minor, resulting from the failure to fulfill the duties incumbent on the parents, pursuant to art. 147 cc."

[160]Cass. Civ., sec. VI, sentence no. 12410/2020.

the possibility of monitoring the machine - unless we think of a remote control system - and would therefore not have the possibility of preventing the harmful event with his conduct.

It would therefore be necessary to go beyond the limits of the "culpa in vigilando" of the preceptor and develop a new concept of "culpa in educando" - which, as previously stated, jurisprudence refers only to parental figures - to be attributed to the person who teaches behavioral paradigms to the AI system.

At the same time, however, it would be necessary to offer this last figure a liberating proof that is not excessively onerous - like that provided for parents by Italian jurisprudence - so as not to discourage the production of systems governed by artificial intelligence.

Assuming that the same programmer has trained hundreds of different machines, he runs the risk of being called upon to compensate for a plurality of crimes committed, suggesting that the single tutor is not the most suitable figure to be held solely responsible also because equipped with limited assets, however, compared to that of the multinational company where it carries out its activity.

The analogical application of the art. 2048 cc, therefore, presents many difficulties in adapting to the hypothesis of liability of a machine, requiring an upheaval in order to be able to translate the key principles and the rationale behind this provision, without however leaving a series of issues unresolved practices.

Consequently, even this type of approach does not seem to be completely convincing.

In conclusion, it is underlined that the provision of the second paragraph of the art. 2048 cc in relation to "masters of art" or profession has lost practical relevance because it remained absorbed by the provision of the art. 2049 cc - also invoked in relation to artificial intelligence liability, as described in the next paragraph - after a regulatory intervention cataloged apprenticeship as an employment relationship.

Art. 2049 cc

The analogical application of the art. 2049 cc for determining the liability of an intelligent machine occupies a position of its own compared to the provisions previously

analyzed.

First of all, the art. 2049 regulates the liability of masters and clients deriving from the illicit conduct of their servants and committed in carrying out the activities for which they are employed.

This provision is peacefully recognized as a hypothesis of objective liability, since the joint and several liability (i.e. it is added to that of the material author of the illicit act) of the owner or client exists due to the mere commission of an illicit act by one of his subordinates, without being offered any type of liberating proof.

In particular, jurisprudence and doctrine identify the following conditions for the liability of the owner or client to arise[161]: i) the existence of a prepositional relationship; ii) harmful (at least) negligent conduct on the part of the servant or clerk; iii) "a causal relationship (rectius, of necessary occasionality) between the damage and the performance of the auxiliary's duties"[162].

To tell the truth, the existence of these requirements opens up a window of opportunity in relation to the possibility for the owner/client to free himself from responsibility: in fact, if he were able to demonstrate the lack of one of them, he could not be held liable for the fait accompli by his subordinate.

As regards the prepositional relationship, it usually exists when it is possible to identify a subject who is employed (typically, an employment relationship) or carries out a specific activity under the direction of another.

Having instead taken into account the second assumption, i.e. the illicit act committed by the person in charge, it must be underlined how this links him closely with the principal, who will have the possibility of making use of all the tools that the clerk/housekeeper would have at his disposal to avoid the onset of liability (for example, by demonstrating the existence of a justifying cause in favor of the material author of the damage).

Finally, as regards the necessary occasional connection required by jurisprudence, it demonstrates the need for even a slight connection between the conduct of the person in charge and the orders received from the principal, in order to ensure that the latter is jointly and severally called for compensation of the damage.

[161]Marco Rossetti, Responsibility of masters and clients, in Commentary on the Civil Code of illicit facts articles. 2044-2059, ed. Ugo Carnevali (UTET legal, 2011), 155.
[162]Cass. Civ., sec. III, sentence no. 6756/2001.

In particular, the necessary occasionality was recognized in the circumstance that "the tasks assigned to the employee made it possible or in any case facilitated the behavior producing damage to the third party."[163]

The jurisprudence takes a fairly rigid position, deeming the connection between the conduct and the orders to exist "even if the employee has operated beyond the limits of his duties, or even transgressing the orders received, as long as it is always within the scope of his duties. "*[164]*

However, the analogical application of the art. has been proposed in doctrine. 2049 cc to resolve the vexed question of the responsibility of the intelligent machine, because apparently it would seem to find some points of contact with the hypotheses formulated.

In particular, the ratio that regulates the rule would seem to support that trend which identifies the allocation of risk as a criterion of non-contractual liability and reserves a mainly compensatory function to such rules.

The provision in question, in fact, regulates the possibility that an auxiliary commits damage in the execution of the orders received from the principal, who must be willing to bear the negative consequences of his subordinate.

In this regard, a part of the doctrine (not only Italian, but also European) has underlined how such a rigid regime of responsibility of owners and clients actually responds to solidarity needs.

The principal, in fact, - generally also identified as the organizer of the business activity (from which he derives the benefits) who is subject to the tasks carried out by the person in charge - has greater assets with which to guarantee the damage compared to the auxiliary and the the absence of exculpatory proof transfers to him the risk of non-compliance of the material author of the damage.[165]

Hence the idea of adopting a similar regime to define the liability regime of the "owner" or in any case of those who benefit from his activity (the same subjects described in paragraph 2.5.1), compared to the "domestic" machine, comparing the human auxiliary to that governed by artificial intelligence technology.[166]

[163]Cass. Civ., sec. III, sentence no. 22058/2017.
[164]Cass. Civ., section III, sentence no. 6632/2008, taking up the words of sentence no. 2574/1999.
[165]Guido Alpa, Civil liability, UTET juridical, 2018, 448.
[166]Maria Costanza, “Robotic Enterprise and Responsibility” in Artificial Intelligence and Responsibility: Responsibility From Algorithm? ; AI and Self-Driving Cars, Production Automation, Medical-

Therefore, upon closer inspection, the applicationanalogue of the art. 2049 of the Civil Code would allow us to avoid the aforementioned problem of the machine's assets, given that the risk is definitively transferred to the owner (natural person owner/user) of the machine.

On the other hand, however, the same difficulty described for the hypothesis of the conferral of "electronic personality" would arise regarding the demonstration of the subjective element of guilt in the behavior of the machine.

In thisThis would exacerbate the already difficult position of the owner/client who would see his already faint hopes of escaping compensation for the damage caused by demonstrating the absence of the application conditions of the art. 2049 cc

In conclusion, although the law seems to embrace the most common hypotheses that could be proposed with respect to an intelligent machine (the domestic or worker robot), it is still necessary to take into consideration the socio-economic aspects determined by such a rigid regulation of responsibility.

It is very easy to imagine that neither an ordinary person nor an entrepreneur would be encouraged to make use of AI systems for domestic or business activities when they encountered a particularly unfavorable regime that did not offer the possibility of escaping compensation for damages possibly caused by them.

This would certainly have negative repercussions on the development and diffusion of intelligent machines, with serious repercussions on their producers and in general on the technological progress of society.

Furthermore, the comparison with art. 2049 cc tends to creak when you ask who actually "gives orders" and consequently generates the "tasks" for which the machines are used.

In fact, it is difficult to identify this person in the user of the machine or in the entrepreneur who uses it in carrying out his activity, as they are not the programmers of these machines.

Consequently, a liability regime that is objectified to the point of making the latter figures liable for the simple fact that the damage was caused by a machine programmed/used (by others) to carry out those specific functions is even more "unfair".

Pharmaceutical Robotization; AI and Contractual Activities; Union Trends and Disciplines, ed. Ugo Ruffolo, Giuffrè Francis Lefebvre, 2017, 112 et seq.

Art. 2052 cc

Finally, all that remains is to analyze the hypothesis feared by some of the doctrine of comparing the damage caused by the intelligent machine to that caused by an animal, always remaining in the perspective of failure to overcome the threshold of subjectivity and treating the AI system as an intelligent res.

As already mentioned, Italian law regulates this eventuality in art. 2052 of the Civil Code, establishing that the owner of an animal is responsible for damage caused by it while it was in his custody or while it was lost/escaped.

Also in this case, the legislator offers the owner of the animal a liberating proof, consisting in the demonstration of the fortuitous event.

Initially, the doctrine had framed this provision among the hypotheses of aggravated liability, i.e. of presumed guilt, which is why the exculpatory proof could consist in the demonstration of the absence of culpable behaviour.

Subsequently, however, this theory had to leave room for what considers art. 2052 cc as a rule of strict liability, currently also shared by the jurisprudence of the Supreme Court of Cassation: "He is liable for the damage caused by an animal pursuant to art. 2052 code civil the owner or whoever has the use of it, for objective liability and not for negligent conduct (even just omission), on the basis of the mere relationship existing with the animal as well as the causal link between the behavior of the latter and the harmful event ."[167]

The release test, consequently, has become more burdensome for the responsible party,[168] having to consist in the general demonstration of an interruptive element of the causal link between the animal's behavior and the harmful event.

In particular, over time, doctrine and jurisprudence have identified hypotheses of fortuitous circumstances capable of eliminating the responsibility of the owner of the animal.

If the latter, in fact, were able to demonstrate that the act of a third party meets the

[167]Cass. Civ., sec. III, sentence no. 17091/2014.
[168]Jurisprudence believes that even the de facto user of the animal can be held accountable. Cass. Civ., sec. III sentence no. 5825/2019: "In terms of damage caused by animals, a distinction must be made, alternatively, between the responsibility of the owner and that of the person who uses it for the time in which it is in use. Keeping the animal in use means exercising over it an effective power of government of the type that normally belongs to the owner, whether this derives from a legal or de facto relationship."

requirements of causal autonomy - that is to say that it was decisive for the verification of the damage, because in its absence the injury would not have occurred - of the unpredictability and exceptionality, would have the possibility of freeing himself from liability pursuant to art. 2052 cc[169]

In the same way, however, the owner of the animal can be exempted or in any case limit his liability when he can demonstrate that the negligent conduct of the injured party, which presents the characteristics listed above, contributed to causing the damage.[170]

If the fault is not so high as to determine the exemption from liability, it may still be taken into consideration for the purposes of determining compensation on the basis of the principles of contributory negligence.

As regards the anomalous behavior of the animal, however, this is recognized by jurisprudence as an internal factor of the animal and as such unsuitable to constitute the owner's exonerating proof, despite its unpredictability.

Now, translating this reasoning to the hypothesis of damage committed by an intelligent machine, a fairly burdensome strict liability regime would emerge for the owner/user of the machine where precisely his only

The basis could be to demonstrate the intervention of an unpredictable, sudden and autonomous external factor that cannot be controlled.

However, yesit is just seen as an internal factor of the animal/machine, as could be the embarassment/malfunction, it is not taken into consideration by the jurisprudence for the purposes of the exculpatory proof.

At most, the owner could demonstrate the fact of a third party (producer/programmer), i.e. the incorrect programming of the machine, to free himself from liability. In this case, however, he would have to provide "technical" proof that is difficult to demonstrate, which could be difficult for a judge to understand and lead to a conviction for failure to achieve the onus probandi.

[169]Guido Alpa, Civil liability, UTET juridical, 2018, 912.

[170]Cass. Civ., sec. III, sentence no. 25223/2015: “The responsibility of the owner or user of the animal for the damage caused by it can be traced back to the hypotheses of presumed responsibility and not to those of presumed fault and is limited only in fortuitous circumstances, i.e. the intervention of an external factor in the determination of the damage, which may also consist of the act of the third party or the fault of the injured party, but which necessarily presents the characteristics of inevitability, unpredictability and absolute exceptionality (in this case, the liability of the riding club for the damage caused by a kick to the face delivered by the horse to an expert rider who had entered the fenced area)"

Furthermore, faced with a malfunction of the machine, one could be tempted - as has been done - to invoke the regulation of the defective product, interrupting the parallelism with the case concerning the animal.

Also in this case the risk of an analogical application of the art. 2052 cc – which has been framed by some of the doctrine as a specification of the art. 2051 cc, in relation to the custody of the thing - consists in the development of an excessively rigid liability regime and without "escape routes" (especially if we consider the provision which refers to the lost or escaped animal), with the consequent socio-economic repercussions described above.

A final reflection, however, deserves to be made in relation to the insurance obligation that was foreseen for the owners of dog breeds - included in a sort of black list drawn up by the Ministry of Health from year to year - deemed to be at risk of causing damage.

This system was then abandoned starting from 2009, driven by some reflections relating to the impossibility of defining a dog as intrinsically dangerous simply by belonging to a breed, to move on to a system of "ex post intervention."

Only if the dog has shown signs of violence can the veterinarian decide to impose the insurance obligation on the owner by registering the animal in a special register.[171]

Both solutions could constitute interesting food for thought: on the one hand, the development of a sort of "black-list" of dangerous machines for which insurance coverage is compulsory could mitigate the burden of an indiscriminate necessity to obtain insurance for very simple AI systems; on the other hand, the possibility of keeping a register of dangerous machines "ex post" could meet those ethical reflections born around the figures of intelligent systems which were previously discussed.

[171]Chiara Cavajoni, Damage caused by animals, in Commentary on the Civil Code of illicit facts articles. 2044-2059, ed. Ugo Carnevali, UTET juridical, 2011, 317.

CHAPTERIII

Artificial Intelligence understood as "good"

After having analyzed the possible extra-contractual liability regimes that could be envisaged in relation to a future machine governed by artificial intelligence - considered effectively intelligent, i.e. capable of making autonomous decisions and escaping the predictability of the programmer - it is appropriate to turn attention to the AI technology available today.

As has been anticipated, to date humanity is not equipped with technological development and computational capacity such as to be able to build machines with such characteristics.

Therefore, the majority opinion within the main bodies of the European Union has taken a position in favor of classifying AI systems as real resources, or rather products.

The objective of this Chapter is to delve into the consequences and repercussions in terms of Aquilian liability deriving from such an approach, using categories of domestic and international law.

In particular, as regards the national legal system, part of the doctrine has developed an analogical application of the art. 2051 of the Civil Code regarding damage caused by things in custody, which, as will be seen, provides for a regime of strict liability borne by the custodian.

Among the various resolutionsof the European Union, however, the hypothesis has been put forward of using the already widespread regulation of liability for defective products, incorporated into the Consumer Code and, if necessary, making the necessary adjustments in order to resolve the critical issues deriving from the simple extension of the same .

Subsequently, the European Parliament signed a proposal for a regulation - a fundamental legal instrument for harmonizing the regulation of artificial intelligence liability in all EU countries - proposing a differentiation of regimes between high-risk systems and low-risk systems, in meaning that will be specified in the paragraph dedicated to it.

Finally, we want to provide a brief reconstruction of the discipline developed by the American system in relation to driverless cars - a reality that is certainly more widespread on US soil than on European soil, despite the first approaches - in order to verify the presence of some useful insights reflection to be translated into the development of the regulation of the damage caused by intelligent machines.

Before delving into these reconstructions, it is necessary to take into account, also as a means of transition between the vision of the AI system as a sentient being and the vision of the AI system as a product, the opinion of a part of doctrine according to which the framing of the the use of artificial intelligence as a "dangerous activity", recalling the precept of the art. 2050 cc

3.1 – Art. 2020 cod. civil "Dangerous activity"

According to the provisions of the art. 2050 cc "Whoever causes damage to others in carrying out a dangerous activity, by its nature or by the nature of the means used, is required to pay compensation, if he does not prove that he has adopted all suitable measures to avoid the damage."

In other words, the rationale of the rule consists in imposing on the person who undertakes a notoriously dangerous activity a more onerous liability regime so that the latter adopts an adequate safety system capable of covering the greatest possible number of harmful events.

Now, since any human activity hides an intrinsic danger, it is appropriate to ask ourselves what can fall into the category of "dangerous activity" according to the jurisprudence.

In particular, the Supreme Court of Cassation had the opportunity to rule on the issue: "dangerous activities are considered not only those classified as such by the public safety law and other special laws, but also those which, by their very nature or the characteristics of the means used entail, due to their marked offensive potential, a significant possibility of damage occurring."[172]

It is therefore necessary to distinguish between typical dangerous activities, identified directly by the legislator - for example the storage and transport of substances at risk of explosion or fire - and atypical ones, the dangerousness of which must be the subject of a concrete assessment by the judge[173], called to distinguish case by case.

In particular, one of the main criteria for determining the dangerousness of an atypical activity consists in evaluating its suitability/probability to cause significant damage.

It is not necessary, as some authors had proposed, to adopt a quantitative criterion, carrying out an ex post evaluation of the damage actually caused in the exercise of such an activity.

Rather, what is important is the possibility of identifying in advance the intrinsic danger of the activity in question or the subsequent danger of the same deriving from the tools

[172]Cass. Civ., sec. III, sentence no. 16052/2015.
[173]Jurisprudence believes that the judgment relating to the dangerousness of the activity that caused the damage cannot be reviewed in terms of legitimacy. See Supreme Court. Civ., sec. III, sentence no. 25028/2019.

used.

The jurisprudence also agrees with this approach, although it refers to the so-called "posthumous prognosis",[174] i.e. an assessment to be carried out after the harmful event, taking into consideration the elements existing at the time the activity was carried out.

Furthermore, as the Supreme Court of Cassation points out,[175]it is necessary to distinguish the dangerousness of the activity - identified precisely in the terms described above of intrinsic suitability to cause damage in relation to the nature or the means used - from the dangerousness of the conduct.

The latter notion in fact refers to an activity which in itself does not present the characteristics of danger, but which takes on its characteristics due to the imprudence or negligence of the operator.

In this case, the negligent conduct of the latter must not be assessed on the basis of the parameters established by the art. 2050 cc, but will rather fall within the general clause of the art. 2043, as it is a constitutive element of the case itself.

Now, the set of considerations just made appears consistent with the letter of the law, because only an activity that can be considered dangerous ex ante can give rise to the obligation on the operator to adopt all suitable measures to avoid any damage.

However, with regard to the probation granted to someone who carries out a dangerous activity, the most authoritative voices of Italian private law doctrine have taken positions that conflict with each other over time.

In fact, since its introduction, the provisions of the art. 2050 cc has given rise to controversy in relation to the possibility of configuring it as a hypothesis of aggravated liability or objective liability, in the sense indicated in the previous chapter.

By carefully reading the Report to the Civil Code, it can be seen that in the mind of the legislator of 1942 this hypothesis should not ignore the subjective element of fault but should constitute "an intermediate solution for which, always maintaining fault as the basis of responsibility, no only the exculpatory proof was placed on the injurer, but the

[174]Cass. Civ., sec. III, sentence no. 19180/2018: "it is therefore necessary to concretely ascertain the requirement of dangerousness with an assessment carried out on a case-by-case basis, keeping in mind that even a non-dangerous activity by nature can become dangerous due to the methods with which it is carried out or the means used to carry it out . The factual investigation must be carried out following the criterion of posthumous prognosis, based on the circumstances existing at the time of the exercise of the activity."

[175]Cass. Civ., sec. III, sentence 8449/2019.

content of the duty of diligence, which is placed on him, was expanded."[176]

Trying to respect this precept, part of the doctrine immediately identified in the rule in question a hypothesis of presumption of guilt, involving an inversion of the burden of proof, falling on the person who carries out the dangerous activity.

Following the approach of this theory, since Article 2050 of the Civil Code regulates a type of liability that is still bound to the subjective element, therefore, the disclaiming proof that allows the operator to be exempt from liability can consist in the demonstration of the absence of fault , proving that suitable measures have been adopted to avoid the damage.

A second school of thought, however, contrasting with the dictates of the law, defines the art. 2050 cc as a hypothesis of objective liability, or a presumption of liability that ignores the subjective element.

In this sense, therefore, the injured party can therefore limit himself to proving the causal link between the harmful event and the carrying out of the dangerous activity, while the operator (injuring party) can free himself from liability only by demonstrating the intervention of an independent external causal factor (i.e. fortuitous event) - be it naturalistic or deriving from human conduct - which, due to its exceptional and unpredictability character, could not be avoided and is therefore suitable for breaking the causal link.

Finally, there is something to point out, for completenessof exposure, a third (minority) position which is seen in the art. 2050 cc a hypothesis of culpa laevissima, which leverages the greater degree of diligence that the operator should comply with in relation to carrying out dangerous activities.

In this regard, the law would require very high levels of meticulousness and attention for the person who deals with the activities in question, compliance with which would therefore constitute the only disclaimer available to the operator to be exempt from liability. Consequently, the injured party would have the possibility of obtaining compensation for the damage suffered even in the presence of a (presumed) very slight fault.

Similarly, over the years the jurisprudence of legitimacy has also oscillated between the

[176]Report to the Civil Code, point n. 795: "On the matter it was not believed that any of the extreme solutions could be adopted: neither one which would attach objective liability to such activities, nor one which would attach ordinary liability for negligence to them."

first two approaches described above.

At first, in fact, the configuration of the art was supported. 2050 cc as a hypothesis of strict liability, with the repercussions that have been examined in terms of exculpatory proof.

In particular, therefore, the Supreme Court of Cassation required the operator to demonstrate the fortuitous event which, "although expressly foreseen as a liberating cause only in the hypothesis referred to in the articles. 2051 and 2052 of the Civil Code, in fact it is relevant in every case of objective liability, on the basis of the general principle that even in these cases of liability, the etiological connection between the generating event and the harmful event is necessary".[177]

However, in recent years the trend has been reversed, resurrecting the category of presumed guilt and aggravated liability, outlining a less onerous regime for the operator in terms of proof of release, at least in principle.[178]

Anyone who carries out a dangerous activity will have the possibility to free themselves from responsibility by demonstrating that they have adopted all suitable measures to avoid the damage caused or that, even if they had adopted them, the damage would have occurred anyway.

The discrimen of the lesser or greater favor of such an approach consists in the more or less literal interpretation of the adjective “all.” In fact, it seems clear that, if one accesses the vision according to which the tortfeasor should adopt "all available measures, even if very new and not yet commonly used, which are suitable for preventing the damage"[179], the release test is in any case particularly burdensome for the operator.

In theory, therefore, the majority interpretation of the jurisprudence outlines a regime that is certainly less burdensome for the injurer compared to a hypothesis of strict liability.

In practice, however, jurisprudence has almost never been satisfied with this demonstration, exempting from liability only in the presence of an external factor that

[177]Cass. Civil section III, sentence no. 8457/2004.
[178]Cass. Civil section III, sentence no. 4590/2020. “to overcome the presumption of guilt placed against him by art. 2050 cc, the simple proof of the unpredictability of the damage is also not relevant, since it must, instead, be demonstrated that it could not have been avoided by adopting the prevention measures that the laws of art or common diligence required."
[179]Andrea Parziale, Art 2050 Ten years lived... dangerously, in Damage and responsibility, 2, 2019, 189-190.

interrupts the causal link, indirectly causing the art. 2050 cc in the context of strict liability.

Now, once the principles and practical implications of the rule in question have been highlighted, it is appropriate to ask whether and how these can be reflected in the field of artificial intelligence technologies.

First of all, it is good to recall the distinction between dangerous activities in themselves and those that become dangerous in relation to the means used.

On the one hand, in fact, it would seem difficult to doubt that the production of AI systems constitutes an intrinsically dangerous activity, since the production of tobacco has also been considered as such.

On the other hand, artificial intelligence technology could constitute the "means used" by which an activity not considered dangerous in itself - for example, the circulation of vehicles - can take on the characteristics of danger, being attracted by discipline of the art. 2050 cc

In the first case, therefore, for the purposes of liability, only the figure of the manufacturer of the machine governed by the AI would come into play, who would therefore be called upon to compensate for the damage caused by such res placed on the market.

One could think of extending this reasoning also to the machine educator, where one wanted to recognize the training and programming of a machine as a dangerous activity, but such a thesis would seem to be a forced widening of the mesh of this last notion and, therefore, not convinces to the core.

The critical issues that emerge from the involvement of only the manufacturer of the AI system have already been highlighted in the previous paragraphs: it would in fact constitute anti-economic behavior and contrary to the principles of business activity, to place on the market a series of products for which being required to compensate the damage caused by them a potentially large number of times.

Even more so when the interpretation provided by the majority jurisprudence regarding the disclaiming proof of the art is taken into consideration. 2050 cc, which in practice, as we have seen, requires an external causal element that interrupts the causal link between the dangerous activity, imposing a regime that almost never leaves the injurer/operator an escape.

The fortuitous case, therefore, could in reality only occur if an external activity was carried out on the machine (hacking of the machine?) or in any case when due to the behavior of a third party - who could also be the owner/user of the same - the machine has been compromised in its functioning.

But also where one wanted to access the vision of the art. 2050 cc as a hypothesis of aggravated liability and the jurisprudence requires the simple demonstration of the adoption of all suitable measures to avoid the damage, the risk could be reversed.

For example, if the scientific community prepared safety standards to be adopted in relation to the production of "intelligent" machines, providing for a whole series of programming techniques or mechanisms for protecting/defending the integrity of the machine, the manufacturer could be exempt from any type of damage by demonstrating that all the provisions imposed have been complied with and the losses suffered would remain with the injured party.

Consequently, once again the issue comes to light which - in the opinion of the writer - is the true key element of the development of a non-contractual liability regime for an artificial intelligence system, i.e. the balancing of the interests at stake: on the one hand, the interest of producers in putting intelligent products on the market, thus contributing to the development and technical progress of human society, without this entailing subjection to a hyper-burdensome liability regime that leaves no room for hypothesis of exemption (this need is actually shared by all potentially responsible subjects); on the other, the injured party's interest in receiving fair compensation for the damage suffered by a machine, regardless of the need to provide "diabolical" or excessively technical proof.

In the second case, where an activity becomes dangerous following the use of AI technology, the responsibility of the "operator" for the damage caused could involve the user - this being the person who uses the dangerous activity - and, possibly , also the manufacturer of the intelligent system on a joint and several basis.

Indeed, the criticisms regarding the jurisprudence's interpretation of the disclaimer pursuant to art. 2050 of the Civil Code can also be extended to this eventuality, even more so if one considers the greater difficulty that the user would encounter in proving the fortuitous event.

Consequently, even the simple analogical application of the art. 2050 cc would not seem

suitable to regulate the liability regime of an intelligent machine, unless it is inserted into a system of checks and balances that ensures a fair balance of interests.

3.2 - Art. 2051 cc

The art. 2051 of the Civil Code regulates "Liability for damage caused by a thing in custody" and establishes that the only way available to the "custodian" to free himself from responsibility is to demonstrate fortuitous circumstances.

The rationale of the rule consists in attributing responsibility for the damage caused (directly) by it to the person who can claim the "custodian" relationship with respect to the thing.

If the damage is not caused by the thing itself, but by the active behavior of the custodian, the latter would be liable on the basis of the general clause of Aquilian liability and, consequently, would be required to pay compensation for the damage only when the damaged party was able to demonstrate the constitutive elements of the case of the art. 2043 cc (described in the previous chapter).

Therefore, first of all it is necessary to focus on what must be considered included in the notion of "things" provided for by the article in question.

Over the years, jurisprudencehas identified a long series of "things" likely to cause damage pursuant to art. 2051 cc, including movable and immovable things, liquids and solids, universal movable goods, inert things and moving things.

In reference to this last category, to tell the truth, it is necessary to take into account a different opinion that has developed in doctrine and jurisprudence, according to which only things that possess an intrinsic dynamism could fall within the provisions of the article in question.

In other words, the distinction between a "dangerous thing" - precisely, that which contains such intrinsic dynamism - and a "non-dangerous" thing, which however can take on its characteristics following the intervention of an external harmful agent, is resurrected.

However, as correctly pointed out by Franzoni in the cited work, this discernment does not deserve to be addressed, since the dangerousness of a thing is not an intrinsic characteristic but something that is potentially assumed by all things when they are lowered within certain factual circumstances.

Consequently, sharing this last approach, all things are likely to give rise to the liability of their custodian, provided that there is a sufficiently high causal link between them and

the harmful event, which, according to ordinary principles, must be demonstrated by the injured party.

On closer inspection, in fact, the dynamism to which the aforementioned opinion refers is nothing other than the causal relationship that exists between the thing and the damage and, in fact, when the judge establishes that there is not sufficient "dynamism" in the thing in order to oblige the custodian to pay compensation, does nothing other than highlight the absence of the etiological contribution of the thing to the occurrence of the event.[180]

Secondly, it is appropriate to try to understand what is the correct meaning to attribute to the custody relationship, in order to have the possibility of determining when a specific person should be considered the custodian of a thing and, consequently, be attributed responsibility for the damage caused by it.

The theorists of responsibility linked to the previously described vision of risk allocation define "custodian" as the person who derives the economic benefits from the exploitation of the thing, the latter having to bear the risk of the damage caused by it.

The most recent jurisprudence opinion, however, has taken a different path in defining the custody relationship, highlighting the so-called "government power" over the thing.

In particular, according to a recent ruling of the Supreme Court of Cassation, “Guardian, therefore, pursuant to art. 2051 cc, is the person who - as mentioned - has "the power to govern" the thing, to be understood as the power to control it, to eliminate dangerous situations that have arisen and to exclude third parties from contact with it."[181]

Therefore, the relevant element for being able to charge responsibility for the damage caused lies in the relationship between the thing and the custodian, the latter being able to be considered as such only where actually, based on the concrete relationship with the thing, he can intervene or in any case take action measures to prevent this from causing damage to other subjects (in this case we speak of the duty of "precaution").[182]

In this sense, the reasoning ofThe Court continues, maintaining that the custody relationship must not necessarily entail specific obligations of control over the thing, but rather postulates "a de facto power over the thing determining the damage, i.e. an

[180]Massimo Franzoni, Illicit Facts: Art. 2043-2059, Rome: Zanichelli, 2020, 410.
[181]Cass. Civ., sec. III, sentence no. 13966/2019, which complies with previous rulings of the Supreme Court: see. Cass. n. 2478 of 2018; n. 15761 of 2016.
[182]Corinna Daini, Liability for things in custody, in Damage and Responsibility, (2, 2019): 193.

effective physical power, which is also combined with legal availability and which implies, therefore, the government and use of the thing itself."[183]

In this regard, the doctrine has correctly established that the power to govern (i.e. the legal and material availability) of the thing can only be assumed by a person who can boast (at least) qualified possession: for example, in the case of the driver of a motionless.[184]

Consistently, jurisprudence has excluded that the mere holder for reasons of courtesy or service - as a subject influenced by the organizational or managerial powers of others

- and the mere user of the thing can be held accountable pursuant to art. 2051 cc, unless, for specific reasons, the grantor has not removed himself from the governance of the same.[185]

Finally, it is necessary to analyze the disclaimer indicated by the article in question, the content of which, as we have already seen, can be interpreted differently depending on the nature attributed to the rule.

Even in this case, in fact, two currents of thought have developed in doctrine and jurisprudence, among those who believe that art. 2051 cc should fall within the hypotheses of aggravated liability and those who instead claim to be in the presence of a hypothesis of objective liability.

As regards the first theory - to be honest, it is a minority one and only partially shared by jurisprudence[186]- the basis of the custodian's liability is found in the violation of the duty to supervise the property.

In particular, there would be a reversal of the burden of proof against the custodian, who would be required to demonstrate either the fulfillment of his obligations by following

[183]Cass. Civ., sec. III, sentence no. 13966/2019.
[184]Cass. Civ., sec. III, sentence n.19657/2014.

[185]Cass. Civ., sec. III, sentence no. 22839/2017: "[...] the power to govern the thing, in the same way as the symptomatic indicators highlighted above (Cass. n. 15779/2006, cit.), cannot be recognized in the hands of anyone who has mere possession of the thing for hospitality or service, operating, in the latter case, within the scope of broader organizational and management powers belonging to others (as already stated in the Court of Cassation, 21 November 1978, n. 5418), or in the hands of those who are merely responsible for the matter user (sporadic or temporary), where the granted right to use the thing does not eliminate, "by specific agreement of the parties, or by the nature of the relationship, or by the factual situation that has arisen", the "power of interference, management and intervention on the thing itself that the grantor has preserved (Cass. n. 15096/2013, cit.)."

[186]Especially in reference to the responsibility of the Public Administration for the maintenance of the roads where the custodian is admitted to prove that the damage could not have been avoided even by exercising the utmost diligence (for example with timely intervention).

the criteria of diligence, prudence and expertise, or that the damage would have occurred equally even if he had done so.[187]

This interpretation, upon closer inspection, would conflict with the letter of the law, where in fact fortuitous circumstances are excluded, which, as explained when speaking of the art. 2050

cc (see previous paragraph), is traditionally referred to hypotheses of presumed liability (not fault).

For this reason, the thesis that frames the art has found much more success in doctrine and jurisprudence. 2051 cc as a hypothesis of objective liability, therefore regardless of any consideration regarding the subjective element.

From this perspective, responsibility must be attributed to the subject for the simple fact that the latter has the power to govern the thing, having the demonstration of fortuitous circumstances as the only exonerating proof available, with his diligent behavior being irrelevant.[188]

Fortuitous chance, as mentioned, represents "that causal factor, extraneous to the subjective sphere and characterized by unpredictability and exceptionality"[189] - may also consist of the conduct of a third party - which has a decisive effect in the occurrence of the harmful event.

Regarding the art. 2051 cc, the doctrine has developed three categories of fortuitous event[190] relevant: i) fortuitous autonomous; ii) accidental accident; iii) fortuitous competitor.

The autonomous fortuitous event (which may consist of a natural fact or the behavior of a third party) becomes such when the external causal factor alone determines the harmful

[187]Recently, again on the subject of PA liability, Cass. Civ., sec. III, sentence no. 18865/2017: "the evidencing proof of the fortuitous event, consisting in the demonstration that the damage occurred in a way that was neither foreseeable nor surmountable with the diligent effort appropriate to the concrete circumstances of the case (in this case, damage caused to the wall structures of an underground room , invaded by water mixed with mud, as a result of the flooding of the neighboring streets during a storm)."

[188]Cass. Civ., sec. III, sentence no. 25018/2020: "The art. 2051 cc, in affirming the responsibility of the custodian of the thing for the damage caused by it, identifies an attribution criterion that ignores any connotation of fault, operating on the objective level of ascertaining the causal relationship between the thing and the harmful event (Cass. 2477/2018). To this end, the conduct of the custodian and the observance of supervisory obligations are not relevant: this responsibility is therefore excluded only by fortuitous circumstances, a factor that does not pertain to the behavior of the person in charge, but to the causal profile of the event (Cass. 15383/2006; Cass. 2563/2007)."

[189]Antonio Scalera, Fortuitous case: the Supreme Court sets things right, Corriere Giuridico, 2, 2019, 213.

[190]Massimo Franzoni, Illicit Facts: Art. 2043-2059, Zanichelli, 2020, 426-427.

event, regardless of the thing or the behavior of the custodian (Franzoni in the work cited gives the example of lightning striking a person who was passing through an avenue surrounded by trees under guard).

The fortuitous event of an accident, on the other hand, occurs when, although the thing in custody has caused the damage, however the etiological determination of the harmful event is attributable to the external causal factor (to remain in the previous example, lightning striking one of the trees "in custody" and the latter falls on the unaware passerby).

The concurrent fortuitous event, on the contrary, is not suitable to absorb the entire causality of the damage but simply constitutes an etiological antecedent, which however cannot be invoked by the custodian for the purposes of exemption from liability (and some even argue for the limitation of itself).

The common element of the fortuitous event is, in any case, the possibility of recognizing the elements of exceptionality and unpredictability, as explained in relation to the provisions of the art. 2050 cc[191]

Furthermore, on the basis of a constitutionally oriented interpretation - still discussed - even the conduct of the injured party who comes into interaction with the thing could be relevant for the purposes of exemption from liability.[192]

In this sense, the duty of precaution mentioned above is contrasted with the duty of caution of those who come into contact with the thing, as the maximum expression of the principle of solidarity established by the art. 2 of the Constitution which permeates our system.

---

[191]A recent ruling of the United Supreme Court (Cass. Civ. sez. un., sentence n. 5422/2021) on the subject of meteorological events speaks of the requirements of exceptionality and unpredictability in these terms: "[…] to be understood, respectively, the first, as the objective improbability of the event and the second as a significant deviation from the normal statistical frequency, capable of making that given event, precisely, an exception."

[192]Cass. Civ., sec. III, sentence no. 8478/2020, recalling a precedent from 2018: "In terms of civil liability for damage from things in custody, the conduct of the injured party, who comes into interaction with the thing, behaves differently depending on the degree of causal impact on the harmful event , in application - even unofficially - of the art. 1227 cc, paragraph 1, requiring an evaluation that takes into account the general duty of reasonable caution, attributable to the principle of solidarity expressed by the art. 2 of the Constitution, so that the more the situation of possible damage can be foreseen and overcome through the adoption by the injured party of the precautions normally expected and foreseeable in relation to the circumstances, the more incidental the causal efficiency of the behavior must be considered imprudent of the same in the causal dynamism of the damage, to the point of making it possible for said behavior to interrupt the etiological link between fact and harmful event, when it is to be excluded that the same behavior constitutes a reasonable or acceptable eventuality according to a probabilistic criterion of causal regularity, characterizing , instead, for the exclusive causal efficiency in the production of the accident."

Finally, before delving into the application of the art. 2051 cc to AI systems, it is interesting here to report the words of a recent decision of the Supreme Court in the field of liability for things in custody which is linked to the balance between the duty of caution and the duty of precaution.

In particular, the Court of Cassation expresses itself as follows: "The vigilance of the custodian, ultimately, comes to be circumscribed by its opposite, that is, by fortuitous circumstances, which translates the general principle of impossibilia nemo tenetur into reference to the position of the custodian. The characteristics of the thing kept, in fact, shape and delimit the fortuitous case, configuring the custodial obligation from an ex ante perspective, i.e. predictability which therefore falls within the legal possibility of fulfilling the obligation itself".[193]

The jurisprudence of legitimacy, in fact, highlights the ancient Latin brocard according to which no one can be required to do impossible things, maintaining that this impossibility must be "calculated" and parameterized also on the basis of the characteristics of the thing guarded.

Once the main characteristics identified by doctrine and jurisprudence in relation to the art. 2051 cc, it is appropriate to verify the opportunity of using this rule to establish the regime of non-contractual liability of artificial intelligence.

The starting point isin assuming that a machine based on AI technology can be described as a res, capable of being included in the scope of application of the art. 2051 cc

In this regard, it is worth mentioning the majority doctrinal opinion according to which the article in question presents a content almost identical to the art. 2052 cc, which regulates liability only for a particular type of res, i.e. animals, although in practice not applied in a perfectly mirrored way.

Secondly, it is appropriate to understand who is the subject held in custody in the sense previously indicated - of the artificial intelligence system that causes damage.

Understandably, the first person to whom the qualification of custodian of the machine is attributed is the owner-user of the robot, a figure already previously mentioned in relation to the other special civil liability regimes described by the Civil Code.

However, some doubts deserve to be raised in relation to the concrete possibility for this

[193]Cass. Civ., sec. III, sentence no. 9693/2020.

last figure to exercise that "governing power" over the machine, understood as the joint material and legal availability of the same.

In particular, one might not be totally convinced by the theory according to which the owner-user is actually able to control the machine, eliminate dangerous situations involving it and exclude contact with third parties.

The risk involved in this case is that of placing a responsibility on a person who is not actually in a position to respect the duty of precaution mentioned above.

Consequently, it is not incorrect to question the possibility of having the role of holder of custody of the machine carried out by a person who adheres more closely to the jurisprudential dictates.

In the first instance, one could refer to the programmer of the machine, but, upon reflection, his figure would lack the de facto availability that is required by the Judicial Authorities.

Rather, the custodian of a machine could be a person in charge of remote supervision, who, having the concrete possibility of preventing the damage from occurring, would be more easily attributable for liability purposes.

However, we are talking about remote hypotheses bordering on science fiction. Already from these first lines the difficulty of an analogical application of the art emerges. 2051 cc to the management of artificial intelligence because the interpretation that has been given to the rule over the years is unable to sufficiently adapt to the peculiarities of the latter.

In any case, even if one wishes to adhere to the thesis that sees the owner-user as the person responsible for the damage caused by the machine, the problem relating to the exonerating proof arises.

Interpreting the art. 2051 cc as a hypothesis of strict liability, the same difficulties encountered during the analysis of the other articles on the subject of exemptions arise.

The demonstration of the fortuitous event in relation to the machine of which one owns or uses, appears particularly burdensome for the responsible party, who therefore risks facing a "no escape" liability, proposing the same economic-social consequences in priorities set out.

Compared to the art. 2050 cc, with which the manufacturer could be involved who by

definition is a more expert person with the machine and who would consequently have greater technical knowledge to be able to find a loophole, the use of art. 2051 cc would potentially place the responsibility on an ordinary individual, who would have greater difficulty in overcoming the burden of proof.

A speechthe opposite could instead be done by interpreting the art. 2051 cc as a hypothesis of aggravated liability and, therefore, taking into consideration the subjective element of fault or more generally, the conduct of the owner-user.

In this way, we would rely on that interpretation which appears to be recurrent if the custodian is a PA, allowing the holder of the duty of custody to demonstrate that the damage would have occurred anyway, despite adopting diligent, prudent and expert behaviour.

In light of the ruling mentioned above which highlighted the principle of a impossibilia nemo tenetur and the parameterization of the fortuitous event in relation to the characteristics of the thing, the owner-user would face an extremely favorable regime, having to demonstrate the simple absence of fault .

Combining this reflection with the consideration made at the beginning relating to the difficulty of the owner-user in concretely exercising that power of government to which the majority jurisprudence refers, it is easy to understand how the burden of proof could be easily overcome, forcing the injured party to bear the losses suffered.

3.3 - Artificial intelligence and defective product liability

Concept of "defective product"

As we have seen, the special civil liability regimes provided for by the Civil Code do not seem to offer an adequate solution to the peculiarities of the intelligent machine as res.

For this reason, quite a few legal experts (Italian and otherwise) have suggested the adoption of an extra-code regime, adopted at community level since the mid-1980s, to attribute responsibility to the producers of AI systems.

In particular, we are referring to defective product liability which was introduced on European soil with Directive 85/374/EEC and implemented into our legal system first with Presidential Decree no. 224/1988, then merged into today's Consumer Code (Legislative Decree 206/2005).

The need for such a discipline began to be felt following the advent of industrial society, through which we witnessed the progressive transition from restricted, artisanal production to mass or series production, typical of large factories.

In this sense, in fact, the rationale of the aforementioned Directive was to protect final consumers - and, in general, buyers of goods produced according to standardized procedures, even for professional or commercial purposes - from damages deriving from a possible "defect" of the themselves, developing a liability regime - harmonized between the various Member States of the European Union - towards the producers who had placed them on the market.[194]

As we will soon see, in fact, the European legislator has provided for a system of liability that the Italian legal system defines as "presumed" (in the EU it is rather called strict liability), based on the simple existence of the link between the harmful event and the product placed on the market.

The idea of using defective product liability to regulate the civil liability of artificial intelligence has been feared since the aforementioned 2017 resolution, where the principle was affirmed according to which, based on the existing legal framework, "the liability by product (according to which the manufacturer of a product is responsible for malfunctions) and the rules governing liability for harmful actions (by virtue of which

[194]Enrico Al Mureden, Product safety and producer responsibility, Giappichelli Editore, 2015, 5.

the user of a product is responsible for behavior that leads to damage) are applicable to damage caused by robots and artificial intelligence".

With the 2017 Resolution, the aim was to lay the foundations for the creation of a unitary liability system at European level, based on a common core - to be launched precisely on the basis of the regulation of defective products - but which had such elasticity as to be able to adapt to technological progress without the need to be constantly updated in relation to the various new products.

However, shortly after, it was the European Commission itself that realized that one of the necessary "next steps" was the adaptation of the key concepts described in the aforementioned Directive to the new models of "products" in circulation and the consequent "adjustment" of the producer responsibility, to be pursued through the creation of a group of experts on the subject as a support body.[195]

The European Parliament also expressed itself in favor of a similar initiative, welcoming the Commission's decision with the Resolution of February 2019, with which it sought to once again urge the need for coordinated regulatory intervention.[196]

After this brief introductory excursus, it is appropriate to focus on the practical aspects of defective product liability - also describing how it was declined by the Italian legislator - in order to highlight the relevant characteristics when directly applied to the damage caused by a machine governed by a technology 'artificial intelligence.

First of all, it is necessary to briefly establish what should be considered a "product", when this can be said to be "defective" and who assumes the role of "producer". In this regard, articles 115 and 117 of the Consumer Code are helpful, which elaborate the articles. 2, 3 and 6 of the Directive in question.

Pursuant to art. 115, in fact, any movable good - including electricity - must be considered a "product", even if it is incorporated into another movable or immovable good.

Over time, there has been an attempt by doctrine and jurisprudence to widen the mesh of

[195]European Commission, Staff Working Document, 25 April 2018, Liability for emerging digital technologies, SWD/2018/137.

[196]Resolution of the European Parliament of 12 February 2019, "A comprehensive European industrial policy on robotics and artificial intelligence", which states that the body: "welcomes the Commission's initiative to create the group of experts on liability and new technologies, with the aim of providing the EU with expertise on the applicability of the Product Liability Directive to traditional products, new technologies and new societal challenges".

this definition in order to include goods that at first sight might not be classified as products: blood used for transfusions, software, a vaccine antihemophilic.[197]

But when can a product be considered defective? The art. 117 provides an answer to this question, listing the fundamental criteria for identifying the defect of a good put on the market.

First of all, the legitimate expectations of consumers in relation to product safety become relevant, which therefore constitute the benchmark for possible defects. They must be parameterized taking into account a series of factors including:

*the)*the perception of safety that consumers could develop towards the product as a result of its promotion/presentation activity, but also on the basis of the instructions and warnings provided by the manufacturer; ii) the (reasonable) intended use of the product and the (foreseeable) behaviors that consumers could adopt in relation to it; iii) the technical knowledge of the time in which it was put into circulation, which is why a product cannot be considered defective simply because a more advanced version of it has arrived and is available on the market.[198]

To these parameters deserve to be added - as correctly also recalled by the jurisprudence of the Supreme Court of Cassation - "other elements that can be assessed and concretely evaluated by the judge of merit, which also include the safety standards possibly imposed by sector regulations ".[199]

In this sense, the subject of product liability intersects with that of product safety, i.e. the policy initiated by the European Union with Directive 83/189/EEC with the aim of providing - through a series of "horizontal" regulatory interventions such as the Directive 2001/95/EC, which establishes the general obligation to introduce safe, or "vertical" products, i.e. which focus on a particular category of products - specific safety standards to be compulsorily respected in production and putting into circulation of the products.

In this regard, it is necessary to recall the art. 103 of the Consumer Code where the definitions of what is generally meant by a "safe" product and a "dangerous" product are offered (the latter definition which can be derived in the negative, i.e. if the good does not comply with the parameters indicated by the standard for the "safe" product then it is

[197]Giuseppe Cassano, Antonio Catricalà, and Renato Clarizia, Competition, Market and Consumer Law, UTET juridical, 2018, 2206.
[198]See art. 119 of the Consumer Code.
[199]Cass. Civ., sec. III, sentence no. 29828/2018.

considered "dangerous").

The notion of a dangerous product, in fact, does not coincide with that of a defective product, since a product may well respect the safety standards required by specific regulations for placing on the market - and be, consequently, safe - but at the same time present a defect resulting from incorrect promotion that has been made or from an error in manufacturing.[200]

Refers toto this last defect, jurisprudence understands it as the error that concerns a single specimen of the series - in coherence with the third paragraph of the art. 117 of the Consumer Code mentioned above - and distinguishes it from the design defect, defined rather as the error referable to the entire series.

As regards the production of goods governed by artificial intelligence, it has been proposed to adopt the safety standards required by Directive 2006/42/EC, better known as the "Machinery Directive", which manufacturers must respect in the development and putting into circulation of "machines".

Continuing further, it is necessary to focus on the figures of the subjects who are called to respond in the event of damage caused by a defect in a product and on the related liability regime.

The producer is defined by the art. 115 paragraph 2-bis of the Consumer Code in these terms: "the manufacturer of the finished product or one of its components [...]". The figures provided for in the art are then added to the final or intermediate manufacturer. 103 lett. d of the Consumer Code: “the manufacturer's representative if the latter is not established in the Community or, if there is no representative established in the Community, the importer of the product; other professional operators in the marketing chain to the extent that their activity may affect the safety characteristics of the products".

This last list includes, for example, "the installer of the product or other professional operators responsible for carrying out safety tests on the products before they are put on the market.[201]

There are therefore a plurality of subjects who can be held jointly and severally liable on

[200]Arianna Fusaro, Defective and dangerous products: responsibilities, in Producer Responsibility, ed. Guido Alpa, Giuffrè Francis Lefebvre, 2019, 368.
[201]Giuseppe Cassano, Antonio Catricalà, and Renato Clarizia, Competition, Market and Consumer Law, UTET juridical, 2018, 2198.

the basis of the legislation in question, where they are held responsible for the same damage. Consequently, the injured party will have the possibility of making full recourse against each of them, and the solvens will have the right of recourse against the others "to the extent determined by the dimensions of the risk attributable to each of them, the gravity of any faults and the entities and the resulting consequences. In cases of doubt, everyone responds equally."[202]

Lastly, it is appropriate to point out the responsibility of the "supplier", which takes on a residual character and intervenes when it is not possible to trace the identity of the producer.

*The ratio*of this responsibility is based on searching for the person with whom the injured person had the last contact and who therefore supplied him with the product in carrying out a commercial activity, because - at least in theory - he is likely to know the information regarding the manufacturer .

Coherently, the art. 116 of the Consumer Code allows the supplier to free himself from responsibility by communicating the identity of the producer or at least of whoever supplied him with the product within three (3) months of the request of the injured party.

Well, as mentioned above, the regulation of liability for defective products requires the injured party to prove the defect, the damage and the causal link between the defect and the damaging event, without taking into consideration the subjective element that can be referred to. to the manufacturer.[203]

Proving the "defect" of a product may not be particularly easy for the injured party, especially when issues of "technical" relevance come into play - for which (expensive) expert advice is necessary - or when the conditions of the product – for example when it is destroyed during or following the damaging event – do not allow easy demonstration.

For this reason, in practice, injured parties tend to resort to proof through simple presumptions, which, according to the law, must be serious, precise and consistent in

[202]Legislative Decree 6 September 2005, n. 206, art. 121.

[203]Summarizes the Cass. setting. Civ., sec. III, sentence no. 29828/2018: "Liability for defective products has a presumed, and not objective, nature, since it does not require the establishment of the manufacturer's guilt, but not the demonstration of the existence of a product defect. It is therefore incumbent on the injured party - pursuant to art. 120 of Legislative Decree no. 206 of 2005 (so-called consumer code), as already provided for by 8 of Presidential Decree 224 of 1988 - proof of the causal connection not between product and damage, but between defect and damage [...]"

order to be admitted by the judge.[204]

It is clear that the greater or lesser probability of success on the part of the injured party in proving the defect of the product to which the damage is attributed depends on the degree of flexibility in the evaluation of these requirements.

For a period, especially in the jurisprudence of merit, judges have tended to favor the position of the injured party, not requiring particularly rigorous proof and often deducing proof of defect directly from the presence of damage.

However, in recent years the Supreme Court of Cassation has tried to remedy this trend and has returned to more rigid positions, requiring a well-argued demonstration.[205]

In this regard, there has been no lack in doctrine of those who have proposed the adoption of different evaluation criteria, of a probabilistic nature, considered more compliant with the peculiarities of a similar discipline: for example, contenting themselves with the impossibility of finding further causes that could have caused the damage.[206]

In particular, we try to leverage that criterion of verisimilitude referred to in the third paragraph of the art. 120 of the Consumer Code regarding technical consultancy, to demonstrate that the discipline does not require real certainty of the defect and that consequently the judge can be satisfied with a simple probability or verisimilitude.

Once the injured party is able to prove the constituent elements indicated above, the manufacturer will have the possibility of being exempt from liability by demonstrating one of the causes of exclusion from liability strictly provided for by the art. 118 of the Consumer Code.

Therefore, the manufacturer will not be held liable for the damage caused by the defective product when he can demonstrate that:

- the product has not been put into circulation, or has not been delivered or shipped to the

[204]Art. 2729 Civil Code.

[205]Lastly, Cass. Civ., sec. III, sentence no. 3258/2016: “Although the proof of the defectiveness of a product can be based on simple presumptions, it does not constitute a correct logical inference to believe that the damage suffered by the user of a product is the unequivocal element of indirect proof of the defective nature of the latter , according to a deductive sequence which, on the assumption of the defectiveness of every product that presents a tendency to cause damage, derives the certainty of the existence of the defect from the mere circumstance that the damage is temporally caused by the use of the product itself."

[206]Arianna Fusaro, “Producer's responsibility: - the difficult proof of the defect”, La Nuova Giurisprudenza Civile Commentata, (6/2017): 898-899.

user purchaser, as specified above;

- the defect that caused the eventharmful arose at a time following its putting into circulation, being able to take advantage of the favorable regime provided for by the second paragraph of the art. 120 of the Consumer Code. In particular, the demonstration of the fact that, taking into account the circumstances, "it is probable" - with reference to the aforementioned criteria other than certainty - that the defect did not exist before the product was placed on the market is considered sufficient;

- the defective product was not manufactured for distribution on the market for consideration or in the exercise of his professional activity;

- the defect is a direct consequence of the adaptation of production safety policies to the provisions of a mandatory legal rule or to a binding provision. The rationale for this exemption consists in not placing the negative consequences of the "error" of the competent Authority on the producer who complied with the instructions given by the latter. The matter, however, takes a different turn if the regulation or measure simply contains a minimum standard of protection, because in that case the manufacturer could well realize the insufficiency of those provisions and adopt "safer" measures.[207]Therefore, simple compliance with minimum safety levels does not in itself exclude the manufacturer's liability;

- the technical and scientific knowledge present at the time the product was put into circulation did not allow the identification of the defect or in any case the qualification of the product as defective (i.e. development risk). The introduction of the exemption in question in the context of defective product liability demonstrates the European legislator's desire to balance two fundamental principles. On the one hand, the manufacturer's duty of precaution, which shares the meaning with the concept previously referred to when talking about the art. 2051 cc, on the other the promotion of progress and technological-scientific development, favored by preventing attributions of responsibility for future risks that were not foreseeable at the time of production which would discourage research. The generic reference "to technical and scientific knowledge" made by the Consumer Code is potentially capable of embracing an immense amount of information and, consequently, the proof that the manufacturer will have to provide will be more or less onerous depending on the criterion adopted to

[207]Andrea Barenghi, Consumer Law, Wolters Kluwer, 2020, 643.

determine the content of his duty to know; the injured party was "aware of the product defect and the danger resulting from it and nevertheless voluntarily exposed himself to it".[208]

Lastly, it is interesting to highlight herethe possibility for the manufacturer or supplier of the component or raw material alone to free themselves from responsibility by demonstrating that the defect derives from the very conception of the product to which the component or raw material belongs. Alternatively, these subjects have the possibility of demonstrating that the latter fully respond to the instructions received from the manufacturer (being considered a sort of "longa manus").

[208]Legislative Decree 6 September 2005, n. 206, art. 122, second paragraph.

3.4 - Applicability of the discipline to Artificial Intelligence Systems

As anticipated, the peculiarities of products governed by artificial intelligence technology have created critical issues regarding the application of product liability in the event of damage caused by them.

In this regard, taking up what was explained in the first Chapter, it appears clear that AI systems constitute something different compared to "traditional" products, because, unlike these, they do not always remain "the same" but evolve - modifying their conduct - depending on the data to which they are exposed.

In fact, it is useful to recall the concept of machine learning to highlight its transformative character[209]of this type of technology and the central role played by learning algorithms, which are truly responsible for the machine's behavior and, therefore, cannot fail to be taken into consideration in the analysis of civil liability (in particular, in relation to the opportunity of involve other subjects than the mere producer).

Proceeding in an orderly manner, we need to ask ourselves whether an "intelligent" machine as a whole can fall within the scope of the definition of "product" provided by Directive 85/374/EEC (taken as seen from art. 115 of the Consumer Code).

In this sense, it is worth mentioning the opinion of some experts according to which the algorithm and the consequent software that animate the AI system should not fall within the notion of "product" that the European legislator had in mind at the time of the issuing of the aforementioned Directive, having to refer to them rather as real digital services.[210]

As such, the liability of the intelligent machine that causes damage should not be regulated by product liability, but rather by Directive 2019/770/EU on digital services, which leaves it up to Member States to develop liability regimes for operators/suppliers of the latter, with the risk of creating a plurality of extremely diversified disciplines in conflict with the need for harmonization.

However, if one really did not want to consider the hardware/software complex of the machine governed by artificial intelligence as a real product, the impasse could be

[209]Giovanni Comandé, Artificial intelligence and responsibility between «liability» and «accountability». The transformative character of AI and the problem of responsibility, in Legal Analysis of the Economy, Studies and discussions on business law, (1/2019), 169-188.

[210]Tiago Sérgio Cabral, Liability and Artificial Intelligence in the EU: Assessing the Adequacy of the Current Product Liability Directive, in Maastricht Journal of European and Comparative Law 27, no. 5 (October 2020), 618 et seq.

overcome by equating the algorithm to a component of the final product, with all the repercussions - illustrated later – on the subject of civil liability.

Therefore, assuming that an AI system can fall within the scope of defective product liability, we must ask ourselves what the defect of such a machine could consist of.

First of all, it is appropriate to recall the aforementioned distinction between product liability and product safety, understanding the latter as the set of sector provisions that must be respected in the development of a specific product for correct placing on the market. .

The application of Directive 85/374/EC to AI systems, in fact, would entail the need to adopt a parallel regulation of this type in relation to production processes, either by extending the parameters described by the aforementioned Machinery Directive, or by developing a new regulation.

Although there are those who argue that compliance with the aforementioned safety standards may be sufficient to eliminate the "defects" of intelligent machines[211], one cannot help but notice that a complete overlap between the notion of safe product and defective product is in contrast with the regulations previously examined and does not correspond to the rationale of the Directive in question.

Indeed, it is necessary to use the tools made available by product liability legislation (i.e. the criterion of reasonable consumer expectation) to identify possible defects.

An overly literal interpretation could lead one to believe that being in the presence of an intelligent product capable of adapting its behavior depending on the circumstances, the consumer could reasonably believe that this could never produce a harmful event (implying that simple causation of the damage, coincides with the defectiveness of the same and considerably worsening the position of the manufacturer).[212]

Therefore, we are inclined towards the idea of being able to identify the defect of the machine, due to the different typologies illustrated above: an AI system may manifest a manufacturing or design defect or may not meet the consumer's reasonable expectations even after an incorrect presentation/warning/instruction thereof by the manufacturer.

In this regard, it is necessary to preliminarily distinguish between: i) the "physical"

[211]Enrico Al Mureden, Autonomous cars and civil liability between current regulations and de iure condendo perspectives, in Contract and Business, (3/2019), 918.
[212]Andrea Amidei, Artificial Intelligence and law - Artificial Intelligence and product liability: developments in European Union law, in Italian Jurisprudence, (2/2019), 1721.

defects of the hardware part (the "dents" to give an idea), for which the applicability of the discipline in question appears undoubted given the affinities compared to traditional cases of defective product liability; ii) defects relating to the software or the learning algorithm, which are more difficult to identify - with the consequences that will be seen in terms of the burden of proof - unless there are glaring errors, such as an oversight in the context of writing the algorithm resulting in an incorrect code string.[213]

If the AI system in question presents ab origine (i.e. at the moment in which it is put into circulation pursuant to art. 119 of the Consumer Code) one of the defects described above, there would appear to be no problems in directly applying the regulation of defective product previously exposed.

The problem exists in relation to the so-called "supervening defect", made possible thanks to the operating mechanism of the intelligent machine.

As mentioned at the beginning of the paragraph, in fact, systems governed by artificial intelligence have the possibility of learning from experience. Consequently, it is not unlikely that a machine is put on the market without defects but that following the collection of data (for example, the habits of its owner) it changes its behavior causing a harmful event.

In light of what has been analyzed previously, therefore, the manufacturer of the machine would not have particular difficulty in demonstrating the non-existence of the defect at the time of putting it into circulation, thus being exempt from liability.

In this way, an extremely unfavorable regime would be created for the consumer/injured party, who would often be forced to bear the losses suffered. Therefore, the Author proposes the adoption of a more elastic notion of "defect" in the context of products governed by artificial intelligence, adapting the legislation to this new type of technology and suggesting the obligation for the manufacturer to implement, in order to to prevent a priori that the machine adopts certain harmful behaviors, "protection blocks", the absence or malfunction of which would lead to the defectiveness of the machine.

In the same vein is the exemption from "development risk" described previously, according to which the manufacturer could free itself from responsibility by claiming that it was not able to know about the defect due to limited technical knowledge at the

[213]Jean-Sébastienne Borghetti, How can Artificial Intelligence be Defective?, in Liability for Artificial Intelligence and the Internet of Things: Münster Colloquia on EU Law and the Digital Economy IV [Internet], eds. Sebastian Lohsse, Reiner Schulze, and Dirk Staudenmayer (Nomos, 2019), 66.

time of production.

It is clear that verifying all the possible behavioral developments that the machine could demonstrate would be impossible to achieve because the manufacturer would face prohibitive costs in terms of money and time, not considering the fact that a significant technical advancement would be necessary.[214]

Consequently, the application of this exemption to the damage caused by AI systems, which as specified can evolve in an unpredictable way, would generate a particularly favorable regime towards the producer, who could benefit from an "ontological" characteristic of intelligent machines to exempt himself from responsibility.

In practice, then, taking the software that is used every day, the so-called bugs are reported by users to the production company, so that the latter releases a new version or a system update to solve the problem, demonstrating that not all " errors" can be predicted in advance.[215]

The Expert Group demonstrated that it has taken this factor into consideration in point 14 of the key findings, where it highlights the need for "the manufacturer to be held responsible for defects in new technologies even if said defects appear after the product has been put into circulation". circulation, as long as the manufacturer is still in control of updates to and updates on the technology".

The last critical issue that deserves to be highlighted is that relating to the difficulty for the injured party to provide proof of the defect in the system governed by artificial intelligence.

Beyond the economic aspect linked to the expensiveness of the inevitable recourse to technical consultancy - however "reversible" on the basis of the provisions of the third paragraph of the art. 120 of the Consumer Code described above - it is particularly burdensome for the injured party to find a way to prove the defect of the intelligent machine - especially when it is attributable to the algorithm - also in light of that jurisprudential orientation which has been mentioned which requires particularly rigorous testing.

In the matter of damage caused by an AI system, it therefore appears more necessary for

[214]Niccolò Filippo Frattesi, Robotics and Algorithmic Responsibility. The production process of artificial intelligence, in Contract and Business, (1/2020), 478 et seq.
[215]Ibid.

the judge's assessment of the existence of the defect to be based on probabilistic or likelihood criteria, rather than certainty.

In this regard, in fact, some authors have suggested the use of alternative methods such as the so-called risk-utility test, consisting in verifying whether the damage could have been avoided by adopting a different algorithm, which at the same time ensures a cost of less or the same implementation for the manufacturer than the one adopted.[216]

[216]Arianna Fusaro, What model of responsibility for advanced robotics? Reflections on the sidelines of the European path, in The New Civil Jurisprudence commented, 6/2020, 1350.

3.5 – The subjects involved

In order to fully understand the development of such a theory, it is necessary to take note of the production process that revolves around the development of an AI system and the consequent involvement of a plurality of subjects.

Three phases can be recognised: i) the implementation phase of the algorithm, which as we have seen is nothing more than the set of rules to follow to arrive at a specific result. The author of the algorithm (probably of multiple algorithms for complex machines like the ones in question), therefore, once the algorithm has been translated into machine language, tests - directly or with the help of other people - his creation in search of of any errors committed; ii) the creation of the software and the copying on computer media; iii) the marketing of the latter, which could consist of selling it to other parties[217] or in sending it to the next link in the production chain, responsible for incorporating it into the hardware part of the AI system (this is referred to as a producer/assembler).[218]

This brief reconstruction, therefore, highlights how it could be reductive to limit responsibility to the simple final producer who puts the AI system into circulation and, consequently, it may be necessary to try to involve additional subjects who work on its creation.

In this sense, the theory is inserted which considers the software a real component of the final product "intelligent machine" and the algorithm a component of the component, making it become a "producer" - as established by the art. 115 of the Consumer Code which incorporates art. 2 of Directive 85/374/EC - also the "manufacturer" of the algorithm.[219]

Consequently, the possibility would be opened up for the injured party to claim compensation also against the latter for the damages caused by an "incorrect evolution" of the machine, making him one of those parties jointly and severally liable - pursuant to the provisions of the art. 121 of the Consumer Code - to the extent determined by the seriousness of any faults and the extent of the consequences.

[217]In this regard, we would like to point out the ongoing debate in the doctrine on the possibility of applying the Directive in question, even if the software was purchased separately from the hardware part, which according to many should be resolved positively in accordance with the rationale of the discipline.

[218]Niccolò Filippo Frattesi, Robotics and Algorithmic Responsibility. The production process of artificial intelligence, in Contract and Business, (1/2020): 458-492.

[219]Ugo Ruffolo, The responsibilities of Artificial Intelligence, algorithm and smart product: for the foundations of a self-learning Artificial Intelligence law, in Artificial Intelligence: Law, Rights, Ethics, eds. Ugo Ruffolo, Guido Alpa and Augusto Barbera, Giuffrè Francis Lefebvre, 2020.

The regulation of the various portions of responsibility would therefore be left to the internal relationships between the producer and the creator of the algorithm - who could be imagined as linked by a contractual bond or a working relationship with the final producer - thus guaranteeing the protection of the consumer-damaged (provided, of course, that he manages to provide the difficult and expensive proof of the algorithm's flaw.)

From this point of view, an interesting food for thought comes from the possibility for the author of the algorithm to make use of the exemption provided for in letter f of the art. 118 of the Consumer Code, proving to have been the long arm of the final producer and to have simply adapted to the instructions received. Indeed, the high technical level involved in the creation and codification of an algorithm would seem to exclude this opportunity, since the manufacturer's directives are unlikely to be so punctual as to exhaust the discretion of the "manufacturer".

Upon closer inspection, bringing these reflections into everyday reality, the involvement of the manufacturer of the component of a machine governed by artificial intelligence technology risks being more theoretical than practical, since the situations in which the final manufacturer and the the author of the algorithm are different subjects are decidedly limited.

In fact, it is necessary to take into consideration that the large IT giants - presumably the (future) major producers of this type of intelligent machines - have developed a tendency to develop "in-house" production processes, i.e. making use of production structures internal for the creation of the various components, with the result that the consumer, if damaged by the defective AI system, can only have recourse against the large manufacturer for compensation for the losses suffered.

Finally, all that remains is to analyze the figure of the trainer (ie trainer) and the position that the latter could possibly occupy within the liability regime thus outlined.

First of all, as underlined during the first chapter regarding machine learning techniques, the data sets that are subjected to a system based on artificial intelligence play a fundamental role for its correct functioning.

In particular, it is the "quality" of the data that has a preponderant importance in the programming field, because the more high-quality data is supplied to the machine, the

higher the level of its performance will be.[220]

Well, the training of the AI system in this sense occurs in a phase prior to its marketing and the person in charge (trainer) may or may not coincide with the author of the algorithm. Consequently, the question arises whether the latter, in the event that it does not provide qualitative data to the intelligent machine, can be held accountable according to the discipline identified by Directive 85/374/EC.

Can the data used to "train the machine to learn" be considered components of the final product? Also in this case, the question takes on greater importance from the point of view of civil liability where the trainer is a third party with respect to the final producer, rather than an "in-house" operator.

According to the opinion of authoritative doctrine, however, the machine trainer would not manufacture "a component" of the final product, but rather should be classified as a service provider, and as such it would not be possible for the injured consumer to act directed towards him on the basis of the Directive in question.[221]

Beyond any contractual liability (depending on the bond that binds them) that the trainer could face towards the end producer, if the person who suffered damage resulting from an AI system wanted to sue the trainer, should use one of the internal categories of non-contractual liability (perhaps art. 2050 regarding dangerous activities? In this case, the trainer should demonstrate that he "has adopted all appropriate measures to avoid the damage", which could include proof of having used high quality data).

After having unraveled the numerous issues raised by the direct application of defective product liability to systems governed by artificial intelligence technology, it is appropriate to make some summary and conclusive considerations.

It seems clear - and the bodies of the European Union have also realized this - that Directive 85/374/EC was not conceived to regulate future intelligent machines, but responded to the needs that developed at the time to offer a solution to the problems due to the prevailing industrial or mass production.

[220]An example of the possible negative consequences of training an AI system with "polluted" data can be found in the Tay Chatbot - created by Microsoft in 2016 - created to deepen communication language. The idea was to create a machine that would develop ever-increasing conversational capabilities by interacting with Twitter users and their posts. A few hours after the launch, Microsoft was forced to withdraw the project because Tay had begun to use offensive and markedly racist expressions.

[221]Andrea Amidei, Artificial intelligence and product liability, in Artificial Intelligence: Law, Rights, Ethics, eds. Ugo Ruffolo, Guido Alpa and Augusto Barbera, Giuffrè Francis Lefebvre, 2020, 136-137.

Consequently, as has been highlighted in the previous paragraphs, critical issues immediately emerged from the attempt to bend the discipline in question to the new challenges posed by AI Systems, dictated mainly by the ability to evolve. In this sense, the main questions concern the possibility of qualifying the machine itself as a product, then the "soul" (i.e. algorithm) of the machine as a component of the same - a theory which has the merit of having tried to involve all relevant parties in the verification of damage due to AI technology[222]-, the need to eliminate the so-called development risk and the problems posed by the "supervening defect" from the examination of exemptions.

Without considering how the main problem of the legislation in question concerns the burden of proof on the injured party, called by a "consumer" - therefore without the technical skills evidently necessary to fully understand the functioning of such machines - to demonstrate the presence of a flaw in the AI System, also incurring enormous technical consultancy costs. In this sense, as is correctly underlined in the doctrine, collaboration on the part of the responsible parties would appear necessary, who should respond to this (for the moment moral) obligation of disclosure and provide information regarding the algorithms, software and data used.

It could then be added that the position of the subject previously described as user/owner who could be the same consumer who purchased the intelligent machine is not taken into consideration, when the latter causes damage to a third party, putting the its possible involvement in the internal categories of non-contractual liability.

Wanting to resolve all the issues raised, therefore, we would encounter upheavals that would distort the regulation of defective products to such an extent - perhaps in contrast with the rationale of the European legislator at the time - that we would begin to wonder whether we are actually using the Directive 85/374/EC as a legal basis from which to start in order to develop an ad hoc regulation for systems governed by artificial intelligence.

Which is not necessarily a bad thing, as long as the need to balance the interests that come into play and recalled when speaking about the art is respected. 2050 cc: on the one hand the interest of the producer (and at this point, also of any creators of

[222]Perhaps the same result could also be achieved by simply including these figures among "the other professional operators in the marketing chain to the extent that their activity could affect the safety characteristics of the products" pursuant to art. 103 of the Consumer Code.

algorithms/trainers) in not being saddled with an excessively rigid or difficult to overcome responsibility and therefore in putting products governed by artificial intelligence on the market , in the name of that scientific progress which lies at the basis of the "development risk"; on the other, the injured party's interest in obtaining compensation for the losses suffered, not having to submit to an excessively technical and/or exorbitant burden of proof.

It is anticipated that, in the opinion of the writer, a measure that should be made in the development of a new liability regime consists in the inversion of the onus probandi to the detriment of the producer/responsible parties, who are certainly in possession of more information which would allow him, by virtue of the traditional principle of proximity of proof, to bear such a burden more easily, having the possibility of easily demonstrating, with data in hand, that the (alleged) defect of the AI system in general or of one of its components in reality is not considered such.

In this sense, the regime that should be outlined will not have to be particularly aggressive towards the producer, who could free himself from responsibility by demonstrating that the verification of that particular harmful outcome would have entailed an effort, from an economic or temporal point of view, which cannot be requested .

When such an eventuality occurs which, it is worth reiterating, is inherent to intelligent machines and constitutes their ontological characteristic, the need for compensation for the damage by the injured party could be satisfied through insurance-type solutions which will be illustrated later in the discussion. .

CHAPTERIV

EU policy in the field of Artificial Intelligence

4.1 – The White Book

During the discussion we had the opportunity to refer to the numerous initiatives undertaken by the bodies of the European Union to try to untangle the various difficulties brought to light by a complex topic such as the regulation of systems governed by artificial intelligence, which undoubtedly represent one of the main challenges of the coming years.

The various efforts made have allowed the European Commission to create the "White Paper on artificial intelligence" - published in February 2020 - which contains the general framework on the topic of intelligent machines and the fundamental principles that have been derived in this regard during the time.

Focusing on the issue of the responsibility of AI systems, the opportunity to evaluate the "possible adjustments of the current EU legislative framework" is further reiterated - with particular reference to the regulations regarding product safety and damage from defective products - always in compliance with the principle according to which “people who have suffered harm caused by the involvement of AI systems must enjoy the same level of protection as people who have suffered harm caused by other technologies; at the same time, technological innovation must continue to develop.”

In this sense, to avoid the creation of excessively burdensome burdens for businesses - in particular for small and medium-sized businesses - the preferable solution according to the Commission would be to adopt an approach based on the concept of risk.[223]

In particular, it would be necessary to identify those "high risk" systems on the basis of two fundamental criteria: i) the application sector, for which it is possible to predict whether there is a significant probability of harmful events; ii) the way in which they are used within the reference sector.

Upon closer inspection, the White Paper does not provide a real definition of "high risk", although the need to adopt clear parameters that are easily understandable and applicable

[223]“White Paper on Artificial Intelligence - A European Approach to Excellence and Trust”, 17-19.

to all interested parties is expressed.

Nonetheless, the idea that emerges from the Commission's paper is to develop an ad hoc regulation for this type of system, i.e. a series of legal provisions applicable only to them, leaving the regulation up to current law - probably to the previously mentioned Directives. other forms of artificial intelligence applications.[224]

The essential elements of the aforementioned new legislation should concern[225]: i) the training data, demonstrating the fundamental role they play in the functioning of intelligent machines, for which safety requirements are suggested regarding the use of data sets large enough to contemplate a high number of possible scenarios ; ii) the mandatory keeping of data and registers for the methodologies used for programming, in order to counteract the "opacity" of the machine, in the sense mentioned previously, i.e. the difficulty in tracing and verifying the decision-making processes of the machine; iii) the information obligations of any subjects responsible towards citizens; iv) the requirements relating to robustness, to be understood as the ability of the machine to resist external attacks of any type, and to precision, to be understood as the probability that the machine will behave according to predictions; v) human surveillance, which responds to the need for an anthropocentric discipline, to ensure which control mechanisms are suggested ex ante, such as the validation of the human being before the machine carries out the action, and ex post, such as the subsequent review of the AI system's decision, or in real time, as a mechanism that allows a human being, in charge of controlling the machine, to deactivate it in the presence of risks; vi) specific requirements for remote biometric identification.

Lastly, it is reiterated that the world of systems governed by artificial intelligence technologies involves a plurality of "operators" potentially attributable to liability, including the developer, the "deployer", i.e. the person who uses a product or service equipped with AI, but also the manufacturer, distributor or importer or private users.

In this sense, the principle that is indicated by the Commission as the foundation of a future liability regime is that according to which the person who is in the best position to face potential risks should be held accountable.[226]From these words emerges the

[224]Ibid, 20.
[225]Ibid., 21 et seq.
[226]Ibid., 25. The Commission's reasoning continues: “For example, if on the one hand AI developers are the most qualified to face the risks deriving from the development phase, on the other their ability to control the risks during use phase may be more limited. In this case the relevant obligation should be

reference to the concept of accountability - already known in terms of personal data protection and GDPR - according to which the person who takes decisions and makes relevant choices must be able to account for them and justify them, responding in case contrary in the appropriate places.

established on the entity applying the AI. This is without prejudice, for the purposes of liability towards end users or other parties who have suffered damage as well as the guarantee of effective access to justice, the question of identifying the person called to answer for any damage caused by the AI system . Under EU law on liability for defective products, this liability is attributed to the manufacturer, without prejudice to national legislation, which may also provide for the possibility of taking action against other parties."

4.2 - The proposed Regulation of 20 October 2020

The fundamental elements of the White Paper just described have actually inspired some initiatives at European level, the most interesting of which culminated with a proposal for a Regulation coming from the European Parliament[227] and intended for the Commission, which attempted to give a concrete form to the indications provided by the latter.

This is a Regulation made up of thirteen articles which bases its discipline - in accordance with what is prescribed by the Commission in the White Paper - on the preliminary distinction between "high risk" systems and "other systems", providing for both 'another two different liability regimes, and on the figure of the operator.

Proceeding in an orderly manner, after an initial declaration regarding the object of the legislation, the first paragraph of the art. 2 of the Regulation establishes that it will apply "in the territory of the Union where a virtual or physical activity, device or process guided by an AI system has caused damage or prejudice to the life, health, physical integrity of a person natural person, to the assets of a natural or legal person or has caused significant non-pecuniary damage resulting in a verifiable economic loss."

The possibility is immediately highlighted that the compensation could concern both the damage caused by death or personal injury, as well as the damage to assets, and - and this is the first difference with defective product liability - damage of a non-pecuniary, provided that it has demonstrable economic value.[228]

The second paragraph of the same article prevents contractual autonomy between private individuals from evading the application of the Regulation, imposing nullity for all contracts that provide for limits to the rights or obligations established therein.

Furthermore, the regulation of the Regulation does not preclude the possibility for the

[227]Resolution of the European Parliament of 20 October 2020, "Civil liability regime for artificial intelligence", the full text of which, including the Regulation under consideration, is available at https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_IT.pdf.

[228]This provision deserves to be read in light of the provisions of Article 5 of the same Regulation, where for high-risk systems a limit of two million euros is established for compensation in the event of death or damage to health/physical integrity and of one million euros for non-pecuniary damages.
Article 6 then adds: "Within the limits of the amount referred to in Article 5, paragraph 1, letter a), the operator held liable in the event of physical damage followed by death of the person concerned shall pay compensation calculated on the basis of costs of medical treatment to which the affected person was subjected before death, and of the pecuniary damage suffered before death due to the cessation or reduction of the earning capacity or increased needs of that person for the duration of the damage before death. The operator held responsible also reimburses the costs of the funeral of the deceased person concerned to the party responsible for covering such expenses."

injured party to bring further liability actions - including the regulation of the defective product, but also of the anti-discrimination, labor and environmental protection legislation - against the responsible party.

Well, the tools necessary for a full understanding of the provisions in question are provided by the art. 3, which focuses on the relevant definitions. In particular, it describes what an artificial intelligence system should be[229], when this must be considered autonomous[230], but above all - and herein lies the great novelty - what the expression "high risk" consists of: "a significant potential in an AI system that operates autonomously to cause damage or harm to one or more people in a random way and which goes beyond what can reasonably be expected; the importance of the potential depends on the interaction between the severity of the possible harm or prejudice, the degree of decision-making autonomy, the probability that the risk will materialize and the method and context of use of the AI system".[231]

More importantly, the art. 3 preliminarily distinguishes two types of "operator" (natural person or legal person, provided that he cannot be considered a "producer")[232]: i) front-end operator, i.e. the person who "exercises a certain degree of control over a risk connected to the operation and functioning of the AI system and who benefits from its functioning", a category which could include both the figure of the owner described in the previous chapter is the deployer, because both could theoretically exercise some sort of control; ii) back-end operator, i.e. the entity that, on an ongoing basis, defines the characteristics of the technology and provides the data and essential back-end support service and therefore also exercises a high degree of control over a risk associated with the operation and functioning of the AI system.

As anticipated, the proposed Regulation outlines two distinct regimes for operators of

[229]Resolution of the European Parliament of 20 October 2020, "Civil liability regime for artificial intelligence", art. 3, "a system based on software or integrated into hardware devices that exhibits behavior that simulates intelligence, including by collecting and processing data, analyzing and interpreting its environment, and taking actions, with a certain degree of autonomy, to achieve specific objectives."

[230]Ibid., "a system based on artificial intelligence that operates by interpreting certain data provided and using a series of predetermined instructions, without being limited to such instructions, although the behavior of the system is linked and aimed at achieving the given objective and other choices operated by the developer during the design phase".

[231]Ibid.
In particular, the Regulation would like to recognize the Commission's power to draw up a list of "high risk" systems to be updated and reviewed periodically.

[232]Ibid. "both front-end and back-end operator, provided that the responsibility of the latter is not already covered by Directive 85/374/EEC"

high-risk AI systems and those of "low-risk" systems: for the former a hypothesis of objective liability arises (art. 4), for the latter instead a regime of liability for presumed negligence (art. 8).

The operator of a system considered to be at high risk will have a single exculpatory proof available, consisting in demonstrating that the damage occurred due to force majeure, with his diligent conduct being of no consequence.

If it is a front-end operator then it will have to take out civil liability insurance, within the limits indicated by the articles. 5 and 6, for AI system operations; similarly, the back-end operator must ensure that its services are covered by an insurance policy that meets the same limits.

The insurance obligation, however, can be considered fulfilled if this is imposed by another national or Union law or there are voluntary company insurance funds that perform the same function.

As regards the operators of other AI systems, art. 8 of the proposed Regulation provides for two causes of exemption from liability, in addition to the already mentioned cause of force majeure: i) the demonstration of the fact that "the AI system was activated without the operator's knowledge and all reasonable and necessary measures to avoid such activation outside the operator's control"; ii) evidence of having used diligent conduct in relation to certain operations such as the selection of an AI system suitable for the task and skills, the commissioning of the system, the monitoring of the system and the maintenance of the same through the installation of all available updates.

It is also appropriate to point out the provisions contained in the third and fourth paragraphs of the same article which respectively provide for the obligation for the operator to pay compensation if the third party who interfered with the AI system - contributing to causing damage - is untraceable or insolvable and the obligation of collaboration with the operator on the part of the manufacturer of the AI system, providing the information necessary to identify responsibility.

The art. 11 then regulates the possibility that the operator can also be the producer pursuant to Directive 85/374/EEC: if the front-end operator is the producer, the regulation of the Regulation prevails over said Directive; if, however, the back-end operator is the manufacturer, then defective product liability must apply to it; in the event that it is possible to link only one operator to the AI system and this is also the

producer, then the provisions of the Regulation should prevail over those of the Directive.

Finally, all that remains is to examine the art. 12 of the Regulation on the subject of recourse action: when there is a plurality of operators, the latter are jointly liable towards the injured party in proportion to their responsibility - to be assessed on the basis of the respective degrees of control - which is why the solvens will have the possibility of taking recourse action against other operators. If it is not possible to obtain from an operator jointly and severally liable the contribution attributable to him, this missing amount is borne by the other operators.

When the operator of a defective AI system fully compensates the injured party, he will have recourse action against the manufacturer of the same on the basis of the regulation of the defective product.

Finally, the operator's insurer who indemnifies the person concerned for damages takes over the latter's position relating to civil liability actions against another person for the same damages, for the amount that was called upon to pay.

It is necessary to recognize how the proposed Regulation just described accepts and tries to decline almost all the principles expressed in the White Paper, especially in relation to the division between high-risk systems, the danger of which lies in the potential to cause damage in a casual manner as well as the limit of predictability, and their periodically updated list.

For this type of system, as seen, a strict liability regime is developed which leaves practically no escape for the operator involved, being able to free himself only due to a cause of force majeure, without however specifying what force majeure could consist of in relation to a AI system.[233]

Such a regime, it is reiterated, may not ensure that balancing of interests to which reference was made above, constituting a very favorable regime towards the injured party and absolutely onerous towards the various operators who may be responsible.

In this sense, both the front-end operator, i.e. the one who can exercise control over the machine - and therefore potentially also the end user who purchases the intelligent machine or the deployer who uses it in carrying out his activity (a doctor who uses this technology to perform a surgical operation), but also other intermediate subjects - both the back-end operator of a high-risk system would have no interest in using, using or monitoring this system, contributing to determining the economic consequences - previously described.

Furthermore, the Regulation does not seem to resolve the critical issues highlighted in terms of defective product liability[234], invoking its direct application when the back-end operator is the manufacturer itself. Only when - and it actually remains to be established which hypotheses we are referring to - the only operator attributable to the AI system that caused the damage is the producer, then the regulation in question will be able to prevail over Directive 85/3734/EEC, eliminating solved the problem.

In this regard, we could suggest greater clarity in identifying the responsible figures and their respective roles, perhaps even dividing them on the basis of their "skills", because it is evident that an "end user" could be less competent than a deployer who uses that technology daily or for which he has received specific training.

[233]An irresistible "external force" could be the hacking of the machine by an external party, but even in this case, if the third party remained untraceable, the operator would still have to respond pursuant to the third paragraph of the art. 8.

[234]As far as the fourth paragraph of the art. 8 provides for the obligation of disclosure of information by the manufacturer, to be understood whether it can also be extended to high-risk systems.

The strong point of the regulation is certainly the provision of the aforementioned insurance obligation, which responds to the needs of compensation for damage within the limits indicated by the art. 5 of the Regulation, however very high.

As regards the regime envisaged for "other types of AI systems", this appears more thoughtful, recognizing a presumed fault on the part of the operator but allowing him to free himself from responsibility with specific exculpatory evidence. The rationale of this provision also seems to respond to that principle of closeness of the burden of proof, because the person with the greatest amount of information available could provide this demonstration more easily than the injured party, without however seriously affecting the position of the person responsible.

Upon closer inspection, however, the use of general expressions such as "reasonable measures" or "due diligence" without giving them a legal meaning could actually create an "interpretative chasm" which risks undermining the correct functioning of the discipline itself. .

In this sense, we could suggest an in-depth study of the necessary reasonable measures and duties of diligence, providing specific contents to these expressions and intersecting the topic with that of product safety (for example by establishing safety standards for placing on the market and a series of specific obligations for subsequent monitoring or maintenance).

In conclusion, the proposed Regulation of October 2020 certainly represents a good starting point for the regulation of civil liability of AI (division of responsibility between different subjects, the compulsory insurance system), but it seems to be open to some criticisms (the onerousness of the strict liability regime, the need to provide greater clarity regarding responsible subjects and exculpatory evidence, the failure to resolve the critical issues of Directive 85/374/EEC), which deserve to be analyzed in the development of legislation ad hoc by the European Union.

4.3 – Artificial Intelligence Act

The European Commission has recently deposited its draft Regulation composed of eighty-five articles entitled "Laying Down Harmonized Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts", with the declared intent of providing a complete and in harmony with the GDPR of legal matters in the field of artificial intelligence. In parallel, a further proposal for a Regulation was presented intended to replace the Machinery Directive, which dictates the new safety parameters - the need for which was underlined in the previous paragraph - for the placing on the market of "machinery products".

At the outset, it deserves to be noted that the proposed Regulation (Artificial Intelligence Act) aims to provide the "general rules" regarding this new type of technology, developing in depth the guidelines expressed in the White Paper, without however making explicit reference to the issue of civil liability - as attempted instead by the European Parliament with the October 2020 proposal - leaving us to predict that the European Union's position on the issue will be postponed to a later date.

In particular, the Regulation aims to: i) harmonize the rules for the introduction on the market, activation and use of AI systems; ii) prohibit certain practices that exploit artificial intelligence; (iii) provide the requirements and duties regarding high-risk systems to be respected by the operators of such systems; iv) dictate harmonized transparency rules for AI systems programmed to interact with natural persons, "emotion recognition" and biometric classification systems, as well as those used to modify video/audio content v) establish the rules for control and surveillance.

The scope of application should concern: providers who put on the market or activate AI systems for users, without revealing their presence on European soil or in a Third State; users of an AI system on the territory of the Union; providers and users of AI systems located in a Third State whose outputs are used on the territory of the Union.

The first major innovation brought by the proposed Regulation - going beyond what has already been examined in the White Paper and the October 2020 Regulation - consists in the distinction between absolutely prohibited AI systems and "high risk" AI systems, for which indications are provided specifications.

It's the art. 5 to identify “prohibited artificial intelligence practices” and to impose a ban on marketing, activating or using: i) AI systems that use subliminal techniques beyond a

person's consciousness to distort material behavior in a way that causes or creates the risk of causing physical or psychological harm to oneself or a third party; (ii) AI systems that exploit the vulnerabilities of a specific group of people related to age, physical or mental disability, to influence the behavior of a person belonging to that group in a way that causes or creates a risk of causing physical harm or psychological to oneself or to a third party; iii) AI systems used by public authorities or for the benefit of the latter, for the assessment or classification of reliability[235] of a natural person in a given period of time, on the basis of his social behavior or his known or expected personal or personality characteristics.

Prohibited activities also include remote biometric identification systems in public places, unless they are used to search for crime victims or search for perpetrators of certain crimes.

Pursuant to art. 6 of the Artificial Intelligence Act, however, an AI system must comply with a double condition in order to be considered high risk: be used as a safety component of a product or be itself a product in accordance with the European regulations indicated in Annex II , and that the latter require a "third party conformity assessment" for the products of which they are components before being placed on the market.

This refers to AI systems used as safety components of medical devices, radio equipment, elevators, toys and the like, for which European legislation provides for a conformity assessment.

There is then a specific list contained in Annex III of the Regulation, updated periodically by the Commission following the parameters indicated by the art. 7, which lists a number of high-risk AI systems in eight areas: 1) post-hoc biometric identification; 2) management of operations relating to main infrastructures (maintenance of public roads or supply of energy, gas, water); 3) professional education and training; 4) employment (e.g. systems that help with personnel selection); 5) access to public or private services; 6) law enforcement; 7) immigration and border control; 8) administration of justice and democratic processes.

The Regulation in question dictates, with articles 9-15, a plurality of measures and requirements that must be adopted in relation to a high-risk system.

[235]The European Union shows that it wants to move away from the social scoring model adopted by China.

First of all, a "risk management system" must be established, implemented and documented, i.e. a "continuous iterative process" to be carried out and updated for the entire life cycle of the AI system, including the identification of all known and foreseeable risks relating to the system, the assessment of the risks that could emerge when the system is used for its "intended purpose" or in the event of any "misuse" or foreseeable risks following the analysis of the data collected by the monitoring systems.

For each AI system it will be necessary to adopt the necessary risk management measures, identified following testing operations to be carried out before it is put on the market. These must consist of measures aimed at eliminating the aforementioned risks, in light of the technical-scientific knowledge available, or at their mitigation when such risks cannot be eliminated, as well as the provision of correct information (if not even actual training) to the end user.

Secondly, a series of provisions are dictated relating to the level of quality of the data with which to train the machine, which must be relevant, representative, error-free and complete, suggesting the development of data governance codes and management practices. Furthermore, technical documentation is made mandatory[236] of the development of the AI system (updated during the life cycle of the machine) to demonstrate compliance with the obligations established by the Artificial Intelligence Act where required and the provision of the AI system with a record-keeping system regarding the operations carried out during its period of activity.

Subsequently, the Regulation prescribes an obligation of transparency and information towards users, who must be able to correctly interpret the machine's outputs and use them correctly, which is why the high-risk AI system should be accompanied by instructions or otherwise concise, complete, correct and clear information regarding the identity of the provider, the characteristics and capabilities of the machine (including its intended purpose, levels of precision and robustness, and so on), the human control exercised, the AI system life expectancy and required maintenance operations.

Another fundamental element recalled by the proposed Regulation is human supervision, which must be possible for the entire duration of the use of the AI system and concerns the prevention or mitigation of the risk of harming fundamental rights. This must be ensured or already under construction when technically possible or it must be identified by the provider before putting the machine on the market and made easily implementable

[236]The content of which is described in detail in Annex IV of the Regulation in question.

by the end user.

Finally, as already explained in the White Paper, AI systems must be designed and implemented in a way that allows the achievement, in light of their intended purpose, of an adequate level of precision, robustness and cybersecurity.

Chapter 3 of Title III of the Regulation in question (articles 16-29) deals with indicating the obligations of the various figures deemed relevant in the context of a high-risk AI system.

Particular attention is paid to the figure of the provider, who is obliged to ensure compliance with the requirements indicated in the articles. 9-15 above and demonstrate the compliance of the AI system when requested to do so by a competent authority. Consistently, the Regulation provides that the provider is the person responsible for keeping the necessary technical documentation, ensuring the compliance procedure described by the art. 43 of the same Regulation and the subsequent posting of the CE mark on the IA system, as well as fulfilling the registration obligation pursuant to art. 51.

The provider then has the obligation to create a "quality management system" - proportional to the size of its organization - promptly documenting it through tools such as policies, procedure protocols or instruction codes, which contain all the relevant elements to verify compliance with the obligations prepared by the Regulation (for example which techniques are used for the design or development of the AI System, which tests or evaluation procedures are carried out on the latter before and after being placed on the market, which monitoring system is used, and so on).

This entity will then be obliged to: i) implement corrective actions - in the most serious cases, withdrawal from the market - when it has reason to believe that the high-risk system attributable to it may no longer be in line with the obligations established by the Regulation; ii) immediately inform the competent authorities if it becomes aware of a risk at a "national level" concerning the health or other fundamental rights of people in the sense indicated by the art. 65 of the Regulation; iii) cooperate with the competent authorities by providing relevant information and documentation upon their request.

The obligations of the provider described so far, pursuant to art. 24 of the Regulation in question, must be fulfilled by the manufacturer if the AI system - in its meaning of software - is put on the market or activated together with a product bearing the brand of said manufacturer. This provision is particularly relevant for the reference made

previously to the IT giants, who will most likely market products governed by an artificial intelligence system and will therefore be called upon to ensure compliance with the Regulation.

Once the analysis of the provider's obligations has been completed, the legislation in question establishes a series of measures that must be respected by various additional figures such as: the representative[237] of the provider, importers and distributors of high-risk AI systems.

In particular, these last two figures - also described in the aforementioned White Paper - share the obligation to ensure that the provider of the AI system has fulfilled all the requirements set out in the Regulation before putting it on the market, informing the competent authorities if a problem may arise. risk from commercial availability. Both will then have to ensure that the compliance of the AI system is not compromised by transport or storage operations, as long as the latter is under their responsibility.

Even the user of a high-risk AI system is not exempt from obligations, which however do not preclude the possibility for national laws to have new ones. The legislation in question, in fact, requires the user to use the AI system by following the instructions provided and on the basis of the latter to carry out regular checks/monitoring on the system itself. If the user has reason to believe that a potentially harmful risk may arise from the use of the AI system, the user is obliged to stop using it and immediately notify the provider or distributor (who must also be contacted if an accident occurs pursuant to art. 62 of the Regulation itself).

Lastly, the art.28 of the proposal in question provides for obligations common to all the figures just analyzed (distributors, importers, users) and to any other third party, who may be considered the providers of the high-risk AI system when one of the following circumstances occurs: i) when market the AI system using their own name or brand; ii) when they change the intended purpose of the high-risk system already put on the market; iii) when they make substantial changes to the AI system.

In fact, if these last two circumstances occur, the "original" provider who put the AI system on the market will be relieved of all obligations under the Regulation in question.

The remaining provisions of the proposed Regulation - which presents notable affinities

[237]Figure who, pursuant to art. 25 of the Regulation, must be identified by the provider established in a third country, and in fact replaces the person represented in the tasks to be carried out on the territory of the Union, such as collaboration with the competent authorities.

with the General Regulation for the protection of personal data (GDPR)[238], which influences the legislation under examination, which is strongly focused on data protection - attempt to create a real "regulatory framework" around this new type of technology.

Firstly, the creation of a Board on artificial intelligence is arranged alongside the European Commission, composed of the European Data Protection Supervisor and the competent national authorities (which will have to be established by the Member States to ensure the application and implementation of the Regulation) , with coordination and support tasks.

The notifying authorities and notified bodies must then be identified, which must operate in an impartial and independent manner, verifying compliance with the various compliance procedures described above while having a plurality of powers available, including the possibility of requesting the presentation of the relevant documentation.

The basic objective consists in the creation of a "network" of authorities which, thanks to the development of common codes of conduct and harmonized procedures at European level, cooperate with each other to achieve common solutions. The collaboration, then, also involves all the subjects relevant to the Regulation, whose already represented obligations to report risks, accidents or malfunctions, activate an information sharing mechanism to ensure the greatest possible safety in terms of artificial intelligence, without give up on technological progress.

In this last regard, in fact, the Regulation itself provides for the possibility of establishing - following specific criteria - some "regulatory sandboxes", particular environments supervised by the European Authorities whose purpose is to ensure the (economic-legal) conditions favorable to the technology development, facilitating feasibility studies or cost-benefit analyses.

Finally, the Regulation, in art. 71, without prejudice to the possibility that, in accordance with the terms and conditions outlined therein, Member States may establish sanctions, including administrative, proportionate, effective and deterrent for the violation of the Regulation, known as the sanctions to be applied to specific transgressions:) an administrative fine of up to €30 million or 6% of global annual turnover - whichever is greater - if the precepts relating to prohibited AI systems or data and data governance for

[238]EU Regulation 2016/679.

systems are not respected high risk; ii) the remaining violations of the obligations established by the Regulation provide for an administrative sanction of up to

20 million euros or equal to 4% of the global annual turnover; iii) the provision of incorrect or misleading information to the competent authorities entails a fine of up to 10 million euros or equal to 2% of the annual turnover.

The amounts to be paid are established on the basis of the nature, severity and duration of the violation and its consequences, on the recidivism of the operator[239] and on the size and market shares of the latter.

[239]A term that includes all the figures examined so far (provider, user, representative, importer and distributor).

4.4 – References to the American system: preemption doctrine and Manufacturer Enterprise Responsibility

Once the need for sectoral regulatory intervention by the European Union bodies to regulate non-contractual liability for damage caused by a system governed by artificial intelligence technology has been highlighted, it could be an interesting food for thought to look overseas and observe the discipline developed on American soil on the subject of a specific AI system: driverless cars.

The starting point of this analysis is certainly the National Traffic and Motor Vehicle Safety Act (1966)[240], American legislation that requires automobile manufacturers to comply with safety standards to protect the public from the risks of accidents caused by the design, construction or operation of automobiles.[241]

In particular, the manufacturer is obligedto: i) put on the market or import into the USA vehicles that comply with the safety standards established by the National Highway Traffic Safety Administration (NHTSA), which succeeded in 1970 the National Highway Safety Bureau (NHSB) which in turn had been established by Highway Safety Act (1966) to give uniformity to the parameters; ii) deliver copies of registers, draw up reports, provide information and allow inspections when requested; iii) issue a certificate of conformity to the safety standards indicated above, when said conformity is contested or questioned; iv) notify the competent Authority of any defects detected.[242]

Well, the American constitutional system, and in particular the so-called "Supremacy Clause" contained in the second Section of the VI article of the American Constitution[243], provides that federal law (Constitution and laws issued on the basis of it) cannot be derogated from the heterogeneous regulatory acts promulgated by the various States.[244]

Federal legislation is reserved (and this consists of the so-called preemption doctrine, i.e. the "pre-emption" on) the development of legislation in matters which, due to their

[240]National Traffic and Motor Vehicle Safety Act (1966)
[241]Anthony D. Brant, "National Traffic and Motor Vehicle Safety Act," Encyclopædia Britannica (Encyclopædia Britannica, inc.), accessed April 19, 2021, https://www.bianca.com/topic/National-Traffic-and-Motor-Vehicle-Safety-Act.
[242]National Traffic and Motor Vehicle Safety Act (1966), Section 108 (a).
[243]Constitution of the United States, art. VI, century. 2: "This Constitution, and the Laws of the United States which shall be made in Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United States, shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, any Thing in the Constitution or Laws of any State to the Contrary notwithstanding."
[244]Guerino D'Ignazio, The Constitutional System of the United States of America, Wolters Kluwer, 2020, 225.

intrinsic characteristics, must be regulated in a uniform manner for all member states, avoiding internal conflicts.

Federal regulations, however, coulddo not totally exclude the regulatory power of each individual Member State, which is often entrusted with the task of regulating certain sectors autonomously, always remaining within the limits set by the harmonized matter.

Regarding the National Traffic and Motor Vehicle Safety Act (1966), although there is an example of a preemption provision,[245]at the same time, a clause is found which establishes that compliance with the safety standards indicated therein does not exempt from liability deriving from common law.[246]

At first glance, these two provisions appear to be in conflict with each other, because on the one hand it seems that state law cannot identify different safety standards, but on the other that it is capable of developing a civil liability regulation that goes beyond what is established by federal law.

A pragmatic solution that has been proposed[247] concerns the possibility of invoking the

implied preemption– shaping what the US Supreme Court said in the Geier v. American Honda Motor Company (2000)[248], which is considered a leading case

on the subject and the subject of great debates.[249]

To summarize this theory, it is argued that a liability regime developed at the state level that derogates from the standards contained in a federal law would conflict with the objective of harmonization of the latter, creating contexts that differ from state to state.

[245]National Traffic and Motor Vehicle Safety Act (1966), Section 103, "Whenever a Federal motor vehicle safety standard established under this subchapter is in effect, no State or political subdivision of a State shall have any authority either to establish, or to continue in effect, with respect to any motor vehicle or item of motor vehicle equipment[,] any safety standard applicable to the same aspect of performance of such vehicle or item of equipment which is not identical to the Federal standard".

[246]Mark A. Geistfeld, "The Regulatory Sweet Spot for Autonomous Vehicles," Wake Forest Law Review (53, 2018): 101, where the following provision is referred to as a "saving clause": "Compliance with any Federal motor vehicle safety standard issued under this title does not exempt any person from any liability under common law".

[247]Ibid.

[248]Geier v. American Honda Motor Co., 529 US 861 (2000).

[249]For further information, see Joseph Mulherin, "Geier v. American Honda Motor Company, Inc. Has the Supreme Court Extended the Pre-emption Doctrine Too Far," Journal of the National Association of Adminsitrative Law Judges 21, no. 1 (Spring 2001): 173-212.
According to some, federal law ensures a minimum safety standard, below which no state law could fall, but this does not exclude the possibility that the latter ensures greater protection, thus without prejudice to the civil liability actions that have been developed in this sense by the Member States.

For this reason, liability regimes that deviate from the provisions of the National Traffic and Motor Vehicle Safety Act (1966) must be considered implicitly "preempted" and therefore give way to federal law.

Consequently, according to this theory, a vehicle manufacturer's compliance with federal safety standards must be understood as a complete defense, exempting it from civil liability (other than that relating to incorrect information regarding the product).[250]

This reasoning, therefore, should be translated into the context of driverless cars, for which the United States has undertaken numerous legislative initiatives involving a plurality of bodies and subjects. Without dwelling specifically on the in-depth analysis of these initiatives, it is enough to imagine that the role of the National Traffic and Motor Vehicle Safety Act (1966) could be taken from the Self Drive Act (2017), where there is a clause very similar to the one previously examined which imposes the prohibition for state laws to deviate from the safety and production standards established therein.[251]

Preemption would therefore imply a particularly favorable regime for the producer - very similar to that described when talking about the art. 2050 cc as a hypothesis of aggravated liability - which could limit itself to placing driverless cars on the market that comply with federal standards to be exempt from liability (unless a hardware or information defect can be identified).[252]

Civil liability, therefore, would fall on the user of the vehicle both when the accident involves third parties and when the driverless car does not cause damage to other subjects other than the user himself, for example by crashing into a wall (this is such a case of "one car collision"[253]).

If on the one hand, therefore, the need to preserve technological progress by stimulating the production of this new type of technology would be perfectly satisfied by this regime which is so favorable towards the producer, on the other the need for compensation for damages is essentially put aside, with consequences that may be difficult to accept.

[250]Ugo Ruffolo and Enrico Al Mureden, "Artificial Intelligence and Law - Autonomous Vehicles and Responsibility in our system and in the US system", Italian Jurisprudence, (2/2019): 1714.
[251]Self Drive Act (2017), sec. 3 (1): "No State or political subdivision of a State may maintain, enforce, prescribe, or continue in effect any law or regulation regarding the design, construction, or performance of highly automated vehicles, automated driving systems, or components of automated driving systems unless such law or regulation is identical to a standard prescribed under this chapter."
[252]Ugo Ruffolo and Enrico Al Mureden, "Artificial Intelligence and Law - Autonomous Vehicles and Responsibility in our system and in the US system", Italian Jurisprudence, (2/2019): 1714.
[253]Ibid.

In this regard, the approach to the issue of level 4-5 American driverless cars is worth mentioning[254] which is defined as "Manufacturer Enterprise Responsibility" and is based on the concept according to which the manufacturer of an automated vehicle, when it places it on the market, is aware of the harmful potential of its products and therefore decides to bear this "risk of business".

Consequently, compensation for the damage caused by the driverless car must be attributed to the vehicle manufacturer, the latter being the person effectively able to predict ex ante the portion of responsibility of the authors of the individual components and adopt internal solutions such as indemnity agreement or price adjustment with such suppliers.[255]

The real innovation proposed by this approach, however, consists in the fact that the sums necessary for compensation for the damage will not be paid directly by the manufacturer of the specific vehicle that caused the damage, but will have to be withdrawn from a fund specifically created and fed by all the manufacturers of this type of technology.[256]

It would be a Fund established by federal legislation that prevails (preemption) over state laws, therefore preventing a varied panorama from forming between the different member states.

It is interesting to understand how such a fund should be fed: in an initial phase, each manufacturer of driverless vehicles should pay a share proportionate to the market share occupied - on the basis of the principle according to which, by the law of large numbers, the more automated cars a subject places on the market, the greater the risk that they cause an accident - to create a starting economic availability.[257]

---

[254]The Society of Automotive Engineers (SAE) International has divided vehicles into 6 levels of automation: i) SAE level 0, at this level only the driver operates; ii) SAE level 1, an automated system sometimes assists the driver in some operations (e.g. braking); iii) SAE level 2, the automated system can carry out certain driving actions on its own, which the driver must always monitor together with the surrounding environment; iv) SAE level 3, the automated system also has the ability to control the surrounding environment, but the driver must be ready to regain control when required; v) SAE level 4, the automated system can control both driving operations and the surrounding environment without the need for driver involvement, but can only do so under certain conditions and in specific environments; vi) SAE level 5, the automated system can perform all driving operations in the same conditions in which a human being could.
"Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," SAE MOBILUS, accessed April 27, 2021, https://saemobilus.sae.org/content/J3016_201806/.

[255]Kenneth S. Abraham and Robert L. Rabin, "Automated Vehicles and Manufacturer Liability for Accidents: A New Legal Regime for a New Era," Virginia Law Review 105 (127/2019): 21.

[256]Ibid.

[257]Ibid., 48.

Subsequently, after sufficient time has elapsed to collect the necessary amount of data, each manufacturer should pay a fee based on the number of accidents in which its products have actually been involved.

Well, the moment an accident occurs involving only type 4 or 5 driverless cars[258] the Fund should cover, within the established limits and regardless of the existence of a defect, all physical damage (bodily injury) - but not property damage - resulting from the operation of the vehicle, unless these are attributable to the owner's negligence in terms of failure to update/maintain the product or the driver himself in the event of failure to take control of the vehicle to carry out delicate maneuvers.[259]

The possibility that the damage arises from negligence on the part of these subjects implies the necessary subscription by the latter of an insurance policy which covers both the damage caused by their vehicle and that suffered by it.[260]

It is already clear from these first lines that the creation of such a fund would bring practical advantages, first of all the possibility of avoiding liability actions by the injured parties which can be costly both in terms of time and money.

In this way, in fact, we would avoid pursuing the path of not always feasible reconstruction of the dynamics of the accident in search of the different "portions" of responsibility of the figures involved in the production process, allowing the injured party to access alternative compensation.[261]

Secondly, compensation for the damage would remain – indirectly – borne by the manufacturer (consistent with a logic of substantial justice, since the owner/driver of a driverless vehicle could have limited control of the same) who could adopt choices economic in order to adapt in advance to this form of "responsibility"[262].

On the one hand, it would be incentivized to invest more in the safety of its products,

[258]Please refer to the work cited for further information on "mixed accidents".

[259]Kenneth S. Abraham and Robert L. Rabin, "Automated Vehicles and Manufacturer Liability for Accidents: A New Legal Regime for a New Era," Virginia Law Review 105 (127/2019): 26.
It is noted that the subjects who could benefit from the Fund thus constituted are: the subjects who are on board the driverless car, pedestrians, cyclists, motorcyclists or other bystanders who suffer physical damage provided that obviously there is a sufficient causal link between the operation of the machine and the damage suffered.

[260]Ibid., 48 et seq. In this sense, a complex regulation is hypothesized which provides for a "conduit" role for the insurer who will be the entity responsible for receiving complaints that may involve the Fund.

[261]Ibid. 33. It is also noted that the creation of a specific Fund would eliminate the ethical problems that could arise: if the automated car "decides" to crash causing physical injuries to the driver in order not to hit a child who throws himself into the middle of the road, could I still claim compensation for the damage? This eventuality would however be covered by the fund established for this purpose.

[262]Ibid., 28 et seq.

comparing the amount of compensation that it would be required to pay if the Fund did not exist and the cost to be borne for the inclusion of features to reduce the probability of accidents.

On the other hand, by knowing ex ante the sums to be advanced for compensation for the damage caused by its products, it will also be able to consider them when setting the market price with a view to sustainable production from an economic point of view.

Naturally, such a system requires some counterweights to avoid possible abuses. Therefore, if the manufacturer had adopted clearly incorrect or erroneous behavior, it could be required to pay punitive damages - already known by American law regarding defective product liability - as a further incentive to adopt adequate conduct.[263]

Fromlastly, it is necessary to take into consideration the possible conduct of both the owner, who would not be able to access the Fund when he modified his automated vehicle in a way that contributes to the cause of the damage, and of a third party who could have interfered with the operation of the car either by means of a "cyberattack" or by conduct that affects physical reality (for example, a third party throwing a stone from the overpass and forcing the vehicle to change its trajectory).[264]In this case, the proponents of "Manufacturer Enterprise Responsibility" argue that the injured party would have the possibility of taking action against the "interfering" third party through ordinary liability actions.

The approach just described, although inserted in a system that follows different principles compared to the European one, turns out to be quite consistent with the American ideal of "race to technological progress" while still favoring the position of entrepreneurs/producers and at the same time to ensure the possibility for injured parties to receive compensation without having to suffer too much.

It is also true that the limitation to simple damage to people but not to property - which should be covered by additional insurance according to the authors of the aforementioned theory - would seem to limit the coverage options to extreme situations such as personal injuries and accidents. irreversible consequences caused by them or death, but it nevertheless represents an important food for thought for the creation of a European regulation on the subject of artificial intelligence liability.

[263]Ibid., 36.
[264]Ibid., 37.

Well, after this brief overview of the main characteristics of the Artificial Intelligence Act, it is necessary to verify whether and how these provisions can have repercussions in the civil liability of AI systems.

First of all, we have already had the opportunity to anticipate that no paragraph of the proposed Regulation makes explicit reference to this topic, giving a glimpse of the crossroads that the European Union is facing. In fact, in light of such a Regulation, it is necessary to choose between referring the regulation of liability to the already existing categories, integrated by the elements described above - possibly with a margin of autonomy on the part of the Member States deducible from the term "the Member States decide the rules on sanctions applicable to violations of this Regulation"[265]- and pursue the development of an ad hoc liability regulation harmonized at European level.

The Artificial Intelligence Act - together with the new proposal on machinery products in relation to security standards - is "limited" to providing an elaboration of the concepts expressed in the White Paper regarding the placing on the market of AI systems with attached administrative sanctions. The fact that this did not specifically deal with liability could mean that the European Union is leaning towards the first alternative, although, to be honest, many expect a "second round", with a separate provision on liability civil.[266]

We must ask ourselves, at this point, whether the provisions examined are sufficient to integrate the already existing liability disciplines, ensuring a balance of the interests at stake. Inevitably, in light of what is established by the Regulation, the choice arises between an "objective" liability regime and a regime that takes into consideration the subjective element of the person responsible.

A central role, as we have seen, is played by the definition of AI system, which refers only to the software part capable of generating outputs and by the deliberately general figure of the provider, made such in order to include all the possible diversified situations of placing on the market or “activation” of the high-risk AI system. In fact, if the AI system refers to a specific product, the manufacturer will assume the obligations

[265]Proposal for a Regulation of The European Parliament and of The Council, “Laying Down Harmonized Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts”, trans. Italian art. 71: “In compliance with the terms and conditions laid down in this Regulation, Member States shall lay down the rules on penalties, including administrative fines, applicable to infringements of this Regulation.”

[266]Lucilla Sioli and Roberto Viola, “EU Commission: 'Here is the True Strength of AI Regulation,'” Digital Agenda (Digital Agenda, April 28, 2021), https://www.agendadigitale.eu/cultura- digital/eu-commission-here-is-the-true-strength-of-the-ai-regulation/.

of the provider and, in the same way, the distributor or importer will do so when they put the AI system on the market under their own name or brand name. The end user could also assume such obligations upon himself, if the specific conditions indicated above occur.

As has been noted, the emerging obligations in relation to a high-risk system concern the risk management system, collection and use of data, technical documentation, transparency obligations towards users, the possibility of ensuring supervision human, monitoring after being put on the market, the precision, robustness and cybersecurity of the machine as well as collaboration with the competent authorities.

Among the internal liability regimes, the indication of all these measures could lead to the adoption of a liability system similar to art. 2050 cc regarding dangerous activities, where "the operator" (which in this case should be the provider) will have the possibility of freeing himself by proving that he has adopted - and in this he would be facilitated by the obligation to keep technical documentation that can prove compliance with the requirements of the Artificial Intelligence Act at the request of the Authorities - all[267] the measures necessary to avoid the damage.

In this regard, it would be necessary for internal jurisprudence to modify the interpretation of such disclaiming evidence because, as explained previously, it usually requires the demonstration of fortuitous circumstances, i.e. an external factor capable of interrupting the causal link. Such an external factor could also consist of the behavior of a third party and, consequently, in such cases it is necessary to ask whether the failure to comply with the obligations established by the Regulation by the additional figures identified therein - distributor, user and importer - can be considered suitable.

However, the numerous similarities between this proposed Regulation and the GDPR could reveal the European legislator's desire to use a single category of responsibility when it comes to these "advanced" activities, based on the already mentioned concept of accountability.

This approach would also reflect the declared programmatic intention of building a regulatory ecosystem that is harmonious and consistent with all the other initiatives of the European Union and, consequently, it would appear to be one of the most likely hypotheses in terms of responsibility.

[267]There would no longer be a need to limit the meaning of this adjective, because all the measures would be indicated by the Regulation itself.

In this regard, the GDPR does not provide precise indications in relation to the methods that the data controller must adopt to comply with the obligations set out therein, but limits itself to dictating general and "outline" requirements, leaving the subjects involved the opportunities to modulate their concrete implementation.

Likewise, the Artificial Intelligence Actwould seem to dictate what the objectives to be achieved in relation to a high-risk AI system are, leaving the organizational aspect to the provider to ensure the achievement of such results and to demonstrate, when requested by the competent authorities, that it has adopted the appropriate measures for this purpose. scope.

It therefore seems reasonable to predict, in terms of artificial intelligence, a liability regime that follows in the footsteps of that provided for by the art. 82 of the GDPR[268] in relation to the illicit processing of personal data or which in any case takes on very similar contours.

In particular, the data controller who violates the obligations deriving from the GDPR will be held liable for compensation for damages[269] etiologically determined by said violation, for the simple fact of having failed to comply with these provisions and noting the subjective element (even in this case, it is an objective responsibility). The General Regulation for the Protection of Personal Data, however, establishes that the data controller is exempted from liability if he can demonstrate that the harmful event is in no way attributable to him.

---

[268]Art. 82 GDPR: "Anyone who suffers material or immaterial damage caused by a violation of this regulation has the right to obtain compensation for the damage from the data controller or data processor.
A data controller involved in the processing is liable for the damage caused by his processing that violates this regulation. A data controller is liable for damage caused by processing only if he has not fulfilled the obligations of this Regulation specifically directed at data controllers or has acted in a manner different from or contrary to the legitimate instructions of the data controller.
The data controller or processor is exempted from liability pursuant to paragraph 2 if he demonstrates that the harmful event is in no way attributable to him.
Where more than one controller or processor or both the controller and the processor are involved in the same processing and are, in accordance with paragraphs 2 and 3, liable for any harm caused by the processing, each controller or data controller is jointly and severally liable for the entire amount of the damage, in order to guarantee effective compensation to the interested party.
Where a controller or processor has paid, in accordance with paragraph 4, full compensation for damages, that controller or processor shall have the right to claim from the other controllers or processors involved therein. treat the part of the compensation corresponding to their share of responsibility for the damage in accordance with the conditions set out in paragraph 2.
Legal actions for the exercise of the right to obtain compensation for damage shall be brought before the competent courts in accordance with the law of the Member State referred to in Article 79(2).."

[269]Damage that must always be proven by the injured party, as correctly reported by the jurisprudence of legitimacy. Cass. Civ., sec. I, sentence no. 1931/2017: "In the event of unlawful processing of personal data […] the non-pecuniary prejudice can never be "in re ipsa", but must be alleged and proven by the plaintiff, under penalty of distorting the functions of the Aquilian responsibility."

Internal jurisprudence, however, has not yet had the opportunity to express a definitive opinion on this last provision, having decided on issues that arose before the entry into force of the GDPR and consequently referring to the previous regulation (Privacy Code[270]), where the provision of the art. was called into question. 2050 cc and the related release proof.

In any case, to date, it is presumable to believe that the non-attributability of the damage can be invoked by the data controller who demonstrates that he has correctly complied

– in light of the principle of accountability - to the obligations imposed by the regulation or that the injury derives from an act of a third party.

Translating this reasoning into the context of the damage caused by a high-risk AI system, therefore, the provider may be called to respond when the injured party demonstrates the damage suffered and the causal link between the latter and the failure to comply of the obligations envisaged. If it is able to demonstrate that, on the contrary, all the provisions of the Artificial Intelligence Act have been respected or that the damage is a consequence of the action of a third party (perhaps the distributor, importer or user, who violate the provisions that concern them) , the provider should be exempt from liability.

In the event that there are multiple providers or multiple subjects (user, distributor or importer) who have failed to comply with the obligations set out therein, it is presumable that they will be called upon to respond jointly and severally to guarantee effective compensation to the interested party and the solvens will have the possibility of claiming compensation based on each person's portion of responsibility.

Nonetheless, the European Commission's failure to include a rule that reflects the art. 82 of the GDPR, suggests that professionals have realized the need, deriving from the peculiarities of systems based on artificial intelligence, to make some adjustments in order to avoid regulatory gaps. Now, it is not known whether this void will be filled by the European Union itself or the Member States will be responsible for proposing solutions, but in any case - in the opinion of the writer - a harmonized intervention at European level would make it possible to better address the difficulties that this complex topic entails.

The main problem, it is reiterated, derives from the ability of the "intelligent" machine to

[270]Legislative Decree 30 June 2003, n. 196 (Privacy Code).

learn and evolve its behaviors based on what has been learned over the long term. Although the proposed Regulation seeks to reduce the risks deriving from this ability by creating a post-market monitoring system that allows the provider to verify whether the AI system remains compliant with European requirements during its use, however in requesting the processing of a risk management system for each high-risk system, the Regulation establishes that the known and foreseeable risks associated with the latter are identified and analysed, implicitly admitting that the provider cannot be required to prevent the unforeseeable.

Similarly, in identifying risk management measures, the maximum possible reduction (where elimination is not permitted) of risks must be ensured through correct design and adequate development, or the implementation of mitigation and control measures in relation to risks that cannot be eliminated, thus admitting their existence.

At this point, we need to ask ourselves what could happen if the damage caused by the machine derives from one of these unpredictable/unavoidable risks, because if one wanted to adhere to the liability system described by the GDPR, the provider would face exemption from responsibility by demonstrating that everything possible has been done to reduce it, in compliance with the provisions of the Regulation.

In this eventuality, the injured party would seem to have to - at least in theory - bear the losses suffered without the possibility of taking recourse against anyone, inevitably eliminating the need for compensation of the damage which must be ensured by the rules of liability.

We therefore suggest, as done in paragraph 3.7, the adoption of an insurance system - the Artificial Intelligence Act does not make any type of reference to an insurance obligation on the provider or other subjects - which guarantees compensation for damage in similar circumstances.

The provider - but in the opinion of the writer, also the user and the other figures identified by the Regulation - should sign a mandatory insurance policy for compensation for damage caused by the AI systems that it puts on the market/active and the premiums paid should contribute to fueling a "Fund for victims of artificial intelligence", which intervenes when the damage caused is not attributable to the provider or other responsible party, because it is beyond foreseeability in terms of costs or time.

This counterweight would ultimately have a significant effect in terms of "trust" towards this new type of technology, towards which the public and large economic operators could be wary precisely because of the difficulty in controlling, monitoring and governing it which has emerged in course of the treatment.

CHAPTER V

The techno-legal norm

5.1 - The concept of Cyberspace

Understood as the virtual space in which users and programs, connected to each other through a telematic network, can move and act for the most diverse purposes, Cyberspace is characterized by an "a-spatial" and "a-temporal" connotation which does not allow no state legislation to ensure effective regulation. On the other hand, the "technical" rules deriving from the "architectural" configuration of the web are able to regulate the conduct of users with a criterion of uniformity, allowing certain actions and inhibiting others on a preventive basis. It is possible that such rules operate in a concurrent manner with state legislation or autonomously. The problem therefore emerges of having to establish whether or not they are endowed with the chrism of "juridicality", so as to be considered binding erga omnes in the same way as a rule of state origin. The question can be answered in the affirmative - at least in our national system - in light of the principle of horizontal subsidiarity, made operational by the recurrence of a general interest in the regulation of cyberspace. Subsidiarity, in fact, will involve private autonomy under a dual and competing profile: on a strictly "horizontal" level, acting as a factor of cooperation between authoritative regulation and social self-regulation, the former being responsible for defining the objectives to be achieved and to the second the concretisation of the same into "technical" rules, the functionality of which will therefore be oriented by the use of instances with a "political" connotation; on a "digital" level, as a method of sharing the individual autonomy of the individual user in the choice of a technical standard presented among a variety of options, which allows him to set the levels of protection of his individual prerogatives, in accordance with his self -subjective determination.[271]

Cyberspace, therefore, is a set of real activities (existential, personal, patrimonial) of the user but in a physically non-perceivable way and in a space where the boundaries can only be attributed to the single system but always guaranteeing freedom of access, use and fruition.

In this space it is possible to foresee an absence of "law" (formally understood) but there

[271]Laghi P., Cyberspace and subsidiarity, ESI, 2015.

cannot be the absence of "right" (understood as "need") of the user.

Cyberspace, therefore, requires regulation.

5.1.1 - Regulatory models

They can be found in doctrine[272]different models of cyberspace regulation:

A-Regulation

Free digital space without regulation.

Central Regulation

Total regulation with unitary global legislation, which however represents only a theory, not feasible as the technological approach is neutral while the legal norm always has an ideological foundation.

Self Regulation

Space regulated through so-called codes of conduct, i.e. self-imposed rules. This approach, rather than providing a regulatory function, aims to respect the interests of the provider, as he himself establishes the code.

Co-Regulation

In this case, the regulation of the space starts from the law and can be accepted by the user, whose rights are thus protected when using the service.

Acceptance by the user makes the rule coercible, guaranteeing the guarantee function typical of the state norm, but using the elasticity of self-regulation.

In this model, the State becomes "one" of the subjects that contribute to regulation, losing its taxing function (it becomes, in fact, a level of regulation).

The co-regulation model respects the principle of subsidiarity, both towards the

[272]Ibid.

Constitution and towards the rules of the European Union, favoring the interests of private individuals by giving them binding relevance also at a regulatory level.

In fact, the model embraces the most modern vision of the sources of law which detach themselves from the classic model of "hierarchy" to embrace "competence", where subsidiarity is not only the ordering of competences but also has the possibility of canceling the result of co-regulation.

The current situation of cyberspace regulation follows the co-regulation model, defined as a multi-phase and multi-model competition between heterogeneous factors, through objectives and guidelines, obligations and responsibilities and high levels of security[273], with priority given to the interests of the person over those of an economic nature.

5.2 - The Lex Informatica

Lex informatica is an expression that refers to technical choices that impose behaviors on web users. Indeed, in the case of information technologies, the relationship between actors, actions and law would unfold through bottom-up, aggregative and community logics typical of the web: lightness of constraints, horizontal links between people, equal dimension, fight for the abolition of Digital divide. The non-territorial spatial dislocation of digital communication transforms society into an organizational model that presents the form of a global network. In contrast with this inclusive, open and interactive potential of the Internet, the other side of the information age emerges characterized by a struggle for the dominion of the Internet, implemented through the management and control of the Internet by a «corporate multi-stakeholder governance", directed by a network of multinational IT and telematic communications companies. The Internet becomes centralized and decentralized at the same time, adapting and repolarizing itself in infinite variation, eluding the territory, structuring immaterial confinements of the "global space" and producing a "regulatory toxic cloud", implemented by a technological architecture (TCP/ IP, DNS, PICS) filtering, labeling and evaluation of web content. From this point of view, the expression coined by Lawrence Lessig that «Code is law» demonstrates how the technological architectures of

[273]Ibid.

the Internet contain self-organizing normative codes and languages that establish and control the rules for access and use of the information available on the Internet.[274]

It can be defined as an attempt at exclusively technical regulation, i.e. one that does not involve legislation, of digital activity.

However, the initial prediction is obviously outdated due to practical problems, such as the identification of the perpetrator of the violation and the actual execution of the sanction, making the legal protection of digital activities impossible.

The concept, however, has been updated over time with the provision of the introduction of legal rules through programming ("by design"): this manages to make infringement impossible in a preventive way (ex ante function), with the prediction which becomes self-effective and no longer requires a controller of the transgression.

The functions of design, i.e. regulation-oriented design, can be summarized as follows:

- *promotional*: favors certain behaviors on the part of the user through incentives;
- *precautionary*: reduces the negative consequences of harmful events in digital activities;[275]
- *repressive*: exclusion of certain conduct by the user through technical inhibition. The behavior, not approved, cannot be continued.

5.2.1 - The Code and the Technique

In the context of the Lex Informatica, the elements considered to affect digital activity are the law, social norms, the market and technology.

The Technique[276]indicates the inclusion of requests, within the planning framework, aimed at protecting values and interests (e.g. fundamental rights) which, by their nature, change from system to system.

That is, it represents the maximum regulatory capacity, as the technique can even exclude violative behaviors a priori, while the Law becomes the programmer's guideline.

Therefore the Technique becomes a fundamental factor: it has disciplinary potential and has within itself the code of conduct which becomes coercible with (declared) knowledge by the

[274]Maestri E., Lex Informatica – Law, person and power in the age of cyberspace, in Law Science Technology, Esi, 2015, n. 4.
[275]Think of the provisions in the field of personal data processing but also of the request for confirmation in IT operations.
[276]Also referred to in doctrine as "architecture", "programming" and "design".

user.

Nonetheless, the criticality of a purely technical system lies in the lack of an external authority (third party) for assessing the sanction.

5.2.2 – The sectors of Lex Informatica

Main sectors of Lex Informatica are:

- *Access to digital content*: it is an expression of political will. It does not fall within the Law, but rather within Technique.
- *Protection of personal data*: it is required by law but is presented to the user who can accept or not, at the cost of not being able to use the service.
- *Copyright protection*: defined by the hyper-control applied, it has a regulatory framework (general provision of the law), a programmatic choice (from which the sanctions derive) as well as an influence on the part of the author of the work who can choose, through the creative commons, the limits of copyright on one's work (example of co-regulation).

5.2.3 - Differences between formal law and Lex informatica.

Formal law:

- has its source in the Legislator;
- the norm is its summary;
- the scope of application is the legal system;
- the evaluation is ex post (the infringement is followed by a sanction);
- the sanction is entrusted to a third party (jurisdictional power).

Lex informatics:

- the source is the programmer;
- the minimum rule is the code;
- the application is extensive;
- the application is ex ante (preventive) therefore self-executing (avoids the conduct);
- since there can be no infringements, there are no sanctions.

In theory, the technical regulation would lend itself as a universal norm and would be perfectly adaptable. The binding nature could arise only through subsidiarity, given the

ultimate aim of protecting the interests and rights of the user, with the transfer of a legally and hierarchically relevant position (ex art. 118, last co., Constitution).

However, it should be noted that, although formally agreed, adhesion to the Lex informatica is necessary (i.e. necessary to proceed with the services), therefore it is difficult to recognize its purely contractual nature (where adhesion would be voluntary).

5.3 - Privacy "by design" and "by default"

The European Regulation for the protection of personal data (GDPR) requires the data controller to adopt adequate technical and organizational measures in order to protect data from unlawful processing.

Article 25, in particular, introduces the principle of privacy by design and privacy by default, an innovative conceptual approach which imposes on companies the obligation to start a project providing, right from the start, the tools and correct settings to protect data. personal data.

The concept of privacy by design dates back to 2010, already present in the USA and Canada and then adopted during the 32nd World Conference of Privacy Guarantors. The definition was coined by Ann Cavoukian, Privacy Commissioner of Ontario. The principles that govern the system are the following:

- prevent, not correct, i.e. the problems must be assessed in the design phase, and the application must prevent the risks from occurring;

- privacy as the default setting (for example, it must not be mandatory to fill in a field of a form whose provision of data is optional);

- privacy incorporated into the project (for example, the use of pseudonymisation or data minimization techniques);

- maximum functionality, in order to respect all needs (rejecting false dichotomies such as more privacy = less security);

- safety throughout the product or service cycle;

- visibility and transparency of the processing, i.e. all operational phases must be

transparent so that data protection can be verified;

- centrality of the user, therefore respect for rights, timely and clear responses to their access requests.

Ultimately, the personal data protection system must place the user at the center, thus obliging the data controller to provide effective protection from a substantial, not just formal, point, i.e. it is not sufficient that the design of the system complies with the norm if the user is not protected.

Privacy by design requires the owner to implement the technical and organizational measures appropriate to the processing, i.e. measures suitable for implementing the principles of personal data protection effectively. When choosing these measures you must take into account:

- the state of the art (the technological advances available on the market);

- implementation costs (proportionality to the risk is required);

- the nature, scope, context and purposes of the processing;

- risks of different probability and severity for the rights and freedoms of natural persons constituted by the processing;

- the weather.

From Article 25 it can be seen that the GDPR approach is centered on risk assessment (risk based approach), as well as other obligations (e.g. notification to national Guarantors), for which companies will have to evaluate the risk inherent to their activities. With this assessment, the extent of responsibility of the data controller or processor is determined, taking into account the nature, scope, context and purposes of the processing, as well as the probability and severity of the risks to the rights and freedoms of users. The GDPR also includes the definition of risk[277]. This risk assessment must be done when the system is designed, therefore before the treatment begins. Clearly, the type of data processed will also have to be taken into account, so in the presence of processing involving minors' data the obligations will have to be more stringent, considering the fact that the risk is greater. Finally, the state of the technology must be taken into account, so the treatment must be adapted over time.

The privacy by default principle provides that by default businesses should only process

[277] Recitals 75 and 76.

personal data to the extent necessary and sufficient for the intended purposes and for the period strictly necessary for those purposes. It is therefore necessary to design the data processing system ensuring that the data collected is not excessive. so that the interested party receives a high level of protection even if he does not take action to limit data collection (e.g. via opt out).

The principle in question obviously affects all aspects of processing, not only the quantity and quality of data, but also the processing period and the people who can access the data.

The introduction of these two principles obviously forces companies to prepare a privacy impact assessment every time the data processing carried out by the product or service should entail high risks for people's privacy or fall into the categories for which PIA is required under the GDPR. The use of software and applications from third-party companies means that the third-party company must develop the risk assessments of the use of the application, which obviously must be designed in compliance with the regulation. Evaluation which must be appropriately documented to the data controller.

5.4 – The techno-legal norm

As we have seen, in the context of the use of Artificial Intelligence the regulatory provisions are of a compensatory nature[278]: essentially, following the damage that occurs as a result of the use of AI, the vicariously responsible party is required to pay compensation to the injured party.

However, as seen also in the field of cyberspace, an exclusively ex post solution appears limiting, given the possibilities of information technology.

It is no coincidence that the cornerstone of the concept of cyber security is not the recovery of the system after the attack (which is however foreseen on a residual basis), but rather avoiding the attack itself by making it impossible.

The presentation of the "techno-legal norm" falls within the modern perspective of the Lex informatica, i.e. the introduction of the rules of law through programming in order to make harmful behavior (more generally, infringements) impossible (ex ante) through rules self-executing, consequently avoiding not only the sanction but also the need for control.

As part of the development of the thesis, it emerged that, as was reasonable to expect, it is

[278]See Chapter II.

not possible to envisage a "universal techno-juridical rule", i.e. a string of code to be inserted into the programming of intelligent agents so that they automatically acquire the ability to take the “legally correct” decision when required.

However, it is possible to trace back to a sort of "programmer code of conduct", i.e. the guidelines that every programmer should follow to obtain the expected result.

These guidelines can be summarized in:

1) *Regulatory evaluation*: establish the machine code guidelines through legal provisions.
2) *Technical phase*:
   - *i)* risk analysis;
   - *ii)* preparation of security measures (security by design);
   - *iii)* preparation of countermeasures (following an attack);
   - *iv)* preparation of checks for the suitability of the system.
3) *Preparation of technical-administrative responsibilities*(obligations based on role).

Recalling what has been said in the context of the Lex informatica, i.e. the Law that becomes the technical guideline, the first phase is to define these guidelines through the existing regulatory provisions: for example, speaking of driverless cars, the regulatory reference cannot deviate from the Code of the Road, which normatively determines the behavior that must be followed while driving. The point is that the machine, whether through a machine learning system (therefore with influence, albeit limited, of the user) much less through a deep learning system (i.e. simulation of a neural network, like the human brain), will never have to evaluate as executable, even if physically possible (or even preferable in terms of functionality and cost-effectiveness), a behavior prohibited by law.

The second phase, defined as "technical", concerns the definition of the risk possibilities in reference to the actual execution of the "wrong" behavior by the machine, whether it derives from a computational error in the machine itself, or whether the command could be " forced” by the user. This is why it is necessary to include security by design solutions, so that incorrect behavior is prevented in advance (ex ante solution). Nonetheless, it is also necessary to provide fault tolerance solutions, typical of IoT and cyber-physical systems, so that the machine can intervene even after having suffered an attack.

Finally, at the end of the technical phase, the final result must not escape the due control of

a third party who can evaluate its suitability for its functions.

Nonetheless, also based on the European provisions in the field of security, which although foreseen in the field of privacy, can be extensively interpreted, as consequences, to the more general areas of liability (in the presence of damage) and accountability (even without verification of the damage) , it is advisable to define in advance the liability regimes attributable to the various agents involved in the construction and programming of the machine. This prediction, by making the scope of responsibilities between programmer, manufacturer, seller and user "clearer", would also facilitate the necessary insurance coverage of intelligent agents, leading to a unitary risk perception in relation to the activities of Artificial Intelligence.

CONCLUSIONS

This paper aims to offer an overview of the reflections that can be carried out on the subject of civil liability of artificial intelligence, initially describing the incredible evolution that has characterized this type of technology since the "Summer Research Project on Artificial Intelligence", organized in Dartmouth College in Hanover in 1955, to the present. In this regard, we had the opportunity to witness the exponential development of "intelligent machines" in step with technological progress and greater computational capabilities, exploited by scholars from all over the world to develop increasingly advanced programming techniques that have made it possible to teach AI systems to perform increasingly complex tasks.

In this sense, an undoubtedly primary role is played by machine learning and its different forms of learning - supervised learning, unsupervised learning and reinforcement learning which are the basis of today's main applications of artificial intelligence, respectively the "anti-spam" filter ” which applies to certain emails, the analysis of user behavior on large platforms such as Amazon and driverless cars - but also from neural networks and deep learning which have allowed the achievement of goals that were unthinkable until recently.

Once we have examined how a modern AI system operates, the need to understand whether it can actually be considered as an intelligent/sentient being or not was highlighted, in order to verify the possible repercussions on the issue of responsibility and involvement of the subjects called to compensate for the damage caused by it. It was noted that the concept of intelligence is liquid and can take on a different meaning depending on the context and the reference perspective, which is why some of the theories proposed over time were analyzed (the Turing test, the metaphor of Searle's Chinese room and Jackson's high level mentalites). For the purposes of civil liability, it was therefore proposed to assign to the concept of intelligence a meaning very similar to that of free will, wanting to consider a machine effectively intelligent only when it is capable of autonomously making decisions that cannot be foreseen by the programmer because daughters of an independent decision-making will. The current level of technology

artificial intelligence is not sufficiently developed for the creation of such an AI system and the positions of modern scientists present a greater or lesser level of skepticism regarding the possibility (but also the opportunity) that machines with such characteristics could be developed in the future. In fact, many are convinced that a machine, no matter how skilled it is in carrying out a certain action, will never be able to understand what it is actually doing, because human intelligence can never be fully reproduced. Others, however, are more optimistic and believe that one day technological progress will advance to a similar level. Without adhering to one or the other approach, the writer limited himself to underlining how the frenetic technological development of recent years has allowed the achievement of many of those results that in the past had been considered impossible to achieve. Therefore, an a priori exclusion of the creation of an AI system actually equipped with real "intelligence", in the sense previously specified, could appear premature and, for this reason, the opportunity was recognized to reflect on the non-contractual liability of a future machine that presents these features.

Regarding a hypothetical future intelligent machine, a plurality of solutions have been proposed. One of the most fascinating consists in the creation of the so-called "electronic personality", that is to say in the attribution of a new genus of legal subjectivity to the AI system, although the European Economic and Social Committee itself considers it an unacceptable moral hazard. From this precept would derive the ownership of rights - in particular those guaranteed at a constitutional level, such as property or civil and political rights - and of duties, from which the obligation for future intelligent machines to pay taxes. From a liability perspective, the attribution of electronic personality would serve as a regulatory expedient aimed at making the AI system directly responsible for the damage caused. However, such a solution, however intriguing, was not fully convincing for a variety of reasons: on the one hand the problems of applying the art. 2043 cc in relation to the subjective element of guilt, given the complexity that characterizes the reconstruction of the machine's decision-making process; on the other hand, the theoretical difficulty in forming an autonomous and separate asset of the AI system. Finally, the final blow is dealt by the following logical reasoning: it is actually the subjects who "educate" the intelligent machine who teach it how to obey the regulatory precepts, which is why they would still be the ones truly responsible for the damage it caused.

The reference to the human figure who "hides" behind the machine also appears

inevitable in the other solutions analyzed for a future truly intelligent machine, whether one wants to consider the AI system like a minor or in any case incapable, using the categories of responsibility of articles 2047 and 2048 cc, whether it is considered to be an intelligent res like an animal (art. 2052 cc) or the slave of ancient Rome.

By equating the AI system with an incapable subject, the supervisor's duty of supervision (identified as the owner-user or user of the same) would come into play pursuant to art. 2047 of the Civil Code, whose regulations do not in any case seem suitable to regulate the hypothesis of the damage caused by this, due to the difficulty for the responsible party to provide the exculpatory proof as interpreted by the jurisprudence, i.e. that the damage would have occurred even if surveillance had been exercised.

Distortingthe links of the art. 2048 of the Civil Code, at the limits of the iure condendo, some experts have attempted to involve the "preceptor" of the actually intelligent machine within the sphere of responsibility, attempting to identify a sort of culpa in educando (attributable, in reality, only to the parents of the minor and not also to tutors), without, however, developing a sufficiently balanced disclaimer - unlike that developed for parents by Italian jurisprudence - as an exemption from liability, leaving a series of practical issues unresolved.

The position of those who wanted to recognize a domestic worker or shop assistant pursuant to art. 2049 cc (provision that does not establish

no proof of release for the owner or the client). We have had the opportunity to underline on the one hand how difficulties emerge in translating some concepts (for example, the prepositional relationship or the performance of an act that is at least negligent on the part of the clerk) in the field of artificial intelligence, on the other hand how such a rigid regime of responsibility for the responsible subject/natural person has socio-economic repercussions in terms of trust towards this type of technology.

Still others have proposed comparing the future intelligent machine and the servus of the Roman system, trying to leverage the concepts of peculium and noxae Datio. The first could allow the resolution of the problem of attributing a separate asset to the AI system - in this regard, the possibility was highlighted for the dominus of the servus of the new Millennium to take out an insurance policy that limits its liability to the maximum established therein -, the second instead could allow a generic fixing of the maximum compensation amount in the intrinsic value of the machine itself, to be parameterized

also in relation to the defects it presents and therefore profiling a possible liability of the manufacturer of the same. The only perplexities of a liability regime modeled on that of Roman law arise from moral-ethical considerations on the concept of "slavery" and fundamental rights, which is why when faced with an "intelligent" system one could lean towards for the subjectification of the same.

Ultimately, it was proposed to use the regime established by art. 2052 cc, equating the future intelligent machine to an animal. In this regard, it was possible to highlight the onerous nature of this approach for the owner of the AI system, since his only evidencing evidence would consist of an external factor interrupting the causal link which is difficult to demonstrate if referring to intelligence technology. artificial. However, it was found useful to reflect on the possibility of adopting a regime similar to the one - now abandoned - on the subject of "dangerous breeds", suggesting the development of a sort of "black list" of intelligent machines and the consequent insurance obligation for owners of the same.

The third chapter of this paper, however, focused on the analysis of the civil liability of already existing AI systems, not yet sufficiently developed

to be considered "intelligent" and to be included in the category of res (although particularly sophisticated).

As a point of conjunction between the machine-subject and the machine-thing we wanted to use art. 2050 cc, referring to artificial intelligence as a sort of "dangerous activity" and attempting to extend the related liability regime, in light of the jurisprudential interpretation of this rule. In particular, the production of this type of technology could be classified as an intrinsically dangerous activity, while the use of artificial intelligence could constitute the "means used" by which an activity not considered dangerous in itself takes on its characteristics. Although this approach does not appear convincing, leveraging the provision in question: in the first case the need to involve the producer of the AI systems is recognised, however the onerousness of the disclaiming proof is reiterated (jurisprudence requires the demonstration of fortuitous circumstances, although the art. 2050 cc refers to "all measures suitable to avoid damage") and the consequent socio-economic repercussions; in the second case, the responsibility of the owner-user should be added to that of the manufacturer.

The involvement of this last figure could be ensured by analogously extending the art.

2051 cc, according to which whoever exercises "governing power" (ie custody) over the thing will be required to compensate for the damage etiologically attributable to it, being able to be exonerated only by proving an external event interrupting the causal link. However, it was possible to highlight the difficulty in framing the owner-user as a subject capable of demonstrating similar power over the AI system - the same goes for the programmer of the same - hypothesizing that a similar role could be played by a a hypothetical remote control figure. Furthermore, neither the interpretation as a hypothesis of strict liability, nor the interpretation as aggravated liability - which jurisprudence uses when the custodian is a PA - would be satisfactory in creating a balanced liability regime that takes into consideration all the interests at stake.

Having exhausted the "internal" categories of responsibility, attention was turned towards European-derived disciplines to regulate the damage caused by the machine-thing. In particular, an attempt was made to verify the applicability of Directive 85/374/EEC on the subject

of "defective product" to products governed by artificial intelligence, analyzing the critical issues that emerged. The essential core of the latter is constituted by the ability of the intelligent machine to evolve its behavior depending on the data to which it is exposed, which cannot be predicted in advance if it does not bear a huge expense in terms of costs and time. Consequently, the exemptions of "supervening defect" and "development risk", if interpreted literally, could ensure an extremely favorable regime for the manufacturer, not in line with the need for compensation of the damage in favor of the injured party. To "teach" an AI system to learn, as seen, a specific algorithm is created and the latter is trained using the aforementioned techniques, by means of data submission. Considering this algorithm as a component of the component (the software, which is incorporated into the product), it would be possible to also consider the creator of the algorithm as producer (and therefore responsible) - the trainer, if it is a different subject, should be called into question using the internal categories of liability, as the training data is not a component of the product - provided that the injured party is able to prove the defect of the same. In this last regard, it is clear that: on the one hand the meaning of "defect" in relation to an intelligent product remains "opaque", on the other hand the injured party must provide extremely technical proof which requires the necessary intervention of an (expensive) technical consultant but also and, above all, the (spontaneous) collaboration of potentially responsible parties.

The critical issues just exposed, not only in terms of defective product liability, but also with regard to the special regimes provided for by the Civil Code, have demonstrated the need to develop an ad hoc regulation - possibly harmonized at European level - on the subject of civil liability of artificial intelligence that takes all these issues into account.

In fact, the European Parliament with a Resolution dating back to October 2020, following numerous initiatives that have followed one another over the years and culminating in the preparation of the "White Paper", has moved in this direction, bringing to the attention of the Commission a proposal for a Regulation on the topic. In particular, this is based on the distinction between high-risk AI systems and other systems, providing for strict liability and aggravated liability regimes. Examining this proposal in depth, we had the opportunity to observe its positive sides - division of responsibility between a plurality of figures, the compulsory insurance system for high-risk systems - and the negative sides - the onerousness of the strict liability regime, the need to provide greater clarity on the subject of responsible parties and evidencing evidence - reaching the conclusion that a more in-depth ad hoc regulation is still necessary.

We have therefore attempted, before providing a personal consideration on the principles that could be followed by the European Union in the development of such a discipline, to draw some food for thought from the American legislation on the subject of driverless cars - underlining its excessive propensity in favor of producers - and by the theory called “Market Enterprise Responsibility”. In reference to the latter, particular attention was paid to the idea of establishing a Fund - powered by the producers of this type of technology - for compensation for damage caused by intelligent machines.

In light of all the above considerations, the writer has attempted to summarize the essential critical issues to be addressed in the development of ad hoc legislation and to provide a personal reconstruction of the extra-contractual liability regime of the intelligence discipline. In particular, it is necessary to pursue the reconciliation of two fundamental interests: on the one hand, that of producers in putting intelligent products on the market - thus contributing to the development and technical progress of human society - without being burdened with a boundless responsibility of escape; on the other, that of the injured party to be compensated for the damage caused by the AI system and not to bear the losses suffered. To do this, it would be necessary to ensure the involvement of (at least) two figures, the manufacturer and the owner-user. The latter's

responsibility should be based on the actual control - the minimum level of which consists of maintenance and updating - that he can exercise over the machine. The manufacturer's liability should be based on the presumption of the defect - for which a univocal definition is necessary - of the machine causing the damage, this person being in possession of information to prove the contrary. The need for compensation for damages, however, could be satisfied

through the creation of a "Fund for the victims of artificial intelligence", fueled by the insurance premiums of the policies compulsorily subscribed by producers and owner-users, which intervenes in the event that the damage is the direct consequence of the behavioral evolution of the system AI that could not be predicted except with reasonable effort in terms of cost and time.

Finally, it is useful to note that the proposed Regulation presented by the Commission on 21 April 2021, during the development of this thesis, provides further food for thought. Although it does not specifically deal with responsibility - and there is discussion as to whether the EU could return to the topic at a later date - it dictates a series of obligations (risk management, post-market monitoring, technical documentation, etc.) for the providers which suggest a future responsibility mechanism centered on the concept of accountability, similarly to the provisions of the art. 82 of the GDPR. The provider, therefore, would have the possibility of being exempted from liability for the damage caused by the AI system attributable to him if he was able to demonstrate compliance with the provisions of the Regulation or an event interrupting the causal link (even the act of a third party).

However, some provisions of the Regulation suggest the awareness of the Union Bodies of the presence of "unpredictable" risks (again in terms of costs and time) due precisely to the ability of the intelligent machine to learn and evolve its behaviors and which could occur even when the provider has respected all the precepts imposed. Therefore, the writer finally wants to renew the hope for a future adoption of an insurance system that can act as a counterweight to this type of eventuality and maximize the need for compensation for the damage suffered by the victims of artificial intelligence.

Precisely the perspective of only compensation for damage, therefore the adoption of the ex post solution with respect to the damage that has already occurred, is limiting in the context of the use of Artificial Intelligence. In fact, an ex ante solution cannot be ruled out, i.e. capable of preventing the damage, which lies within the programming scope of

the machine.

The proposed "techno-juridical rule", to be understood as a logical operation of convergence between technical rules (specific to the machine), ethical principles (based on behaviour) and legal rules, aims to implement a solution already in the planning phase of the machine which, taking place in compliance with the founding cornerstones of the techno-legal norm, leads the AI to correctly evaluate the decision to be made on tracks not dissimilar to the security by design provisions adopted in the context of the regulation of cyber space.

BIBLIOGRAPHY

AL MUREDEN E., *La sicurezza dei prodotti e la responsabilità del produttore*, Giappichelli Editore, 2015.

ALPA, G. *Diritto e intelligenza artificiale: Profili generali – Soggetti – Contratti – Responsabilità civile – Diritto bancario e finanziario – Processo civile [Internet],* Pacini Editore, 2020.

ALPA, G., *La responsabilità civile*, UTET giuridica, 2018.

ALPA G., *La Responsabilità Del Produttore*, Giuffrè Francis Lefebvre, Milano, 2019.

AMIDEI A. *Robotica Intelligente e Responsabilità: profili e prospettive evolutive del quadro normativo europeo*, in *Intelligenza Artificiale e Responsabilità: Responsabilità Da Algoritmo?; A.I. e Automobili Self-Driving, Automazione Produttiva, Robotizzazione Medico-Farmaceutica ; A.I. e Attività Contrattuali*; *Le Tendenze e Discipline Unionali*, Giuffrè Francis Lefebvre, 2017.

ASIMOV I., *Io, Robot*, New York, 1950.

BARCELLONA M., *Trattato della responsabilità civile*, UTET giuridica, 2011.

BARENGHI A., *Diritto Dei Consumatori*, Wolters Kluwer, Assago, 2020.

BENEVENTANO D. - BERGAMASCHI S. - Claudio SARTORI C., *Fondamenti diInformatica*. Società Editrice Esculapio, 2020.

BIANCA, Massimo Cesare. *Diritto civile 5. La responsabilità.* Giuffrè Francis Lefebvre, 2021.

BOBBIT P. – CALABRESI G., *The tragic choices*, New York, 1978.

BONACCORSO G., *Mastering Machine Learning Algorithms: Expert Techniques for Implementing Popular Machine Learning Algorithms, Fine-Tuning Your Models, and Understanding How They Work*, Packt Publishing, 2020.

BORRUSO R. - RUSSO S.- TIBERI C., *L'informatica per il giurista*. GiuffrèEditore, 2009.

BRUNO E. – CRICENTI G. – PARISI I., *Diritto della responsabilità civile: Lezioni*, Pacini Editore, 2016.

BUFFONE G., *Responsabilità civile automobilistica*, CEDAM, 2016.

CALABRESI G., *The Costs of Accidents*, Yale University Press, 1970 (tradotto in *Costo degli incidenti e responsabilità civile*, Giuffré, 1970).

CARINGELLA F. - BUFFONI L., *Manuale Di Diritto Civile*, Dike giuridica,2015.

CASSANO G. – CATRICALÀ A. – CLARIZIA R., *Concorrenza, Mercato e Diritto Dei Consumatori*, UTET giuridica, Torino, 2018.

CASTRONOVO C., *La Nuova Responsabilità Civile,* Giuffrè, Milano, 2006.

CAVAJONI C., *Danno cagionato da animali* in *Commentario del Codice Civile dei fatti illeciti artt. 2044-2059*, UTET, 2011.

CENDON P., *Responsabilità Civile*, UTET giuridica, Torino, 2020.

CERINI D., *Il diritto e gli animali: note gius-privatistiche*, Giappichelli Editore, Torino, 2012.

CIACCI G. – BUONUOMO G., *Profili di informatica Giuridica*, Wolters Kluwer Italia S.r.l., 2018.

CIRILLO G.P., *Sistema Istituzionale Di Diritto Comune*. Seconda ed. Assago: Wolters Kluwer, 2021.

COLUCCI D'AMATO L. - DI PORZIO U. *Introduzione Alla Neurobiologia: Meccanismi Di Sviluppo, Funzione e Malattia Del Sistema Nervoso Centrale*, Milano, Springer, 2011.

COMPORTI M., *Fatti illeciti: le responsabilità oggettive*, Giuffrè Editore, 2009.

COPPIN B., *Artificial Intelligence Illuminated*, Jones and Bartlett Publishers, Boston, 2004.

COSTANZA M., *Impresa Robotizzata e Responsabilità*, in *Intelligenza Artificiale e Responsabilità: Responsabilità Da Algoritmo? ; A.I. e Automobili Self-Driving, Automazione Produttiva, Robotizzazione Medico-Farmaceutica ; A.I. e Attività Contrattuali ; Le Tendenze e Discipline Unionali*. Ed. Ugo Ruffolo, Giuffrè Francis

Lefebvre, 2017.

D'ACQUISTO G. - NALDI M. - BIFULCO R. - POLLICINO O. - BASSINI M., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli Editore, 2018.

DE SPIEGELEIRE S. - MAAS M. - SWEIJS T., *Artificial Intelligence And The Future Of Defense: Strategic Implications For Small- And Medium-Sized Force Providers*, Hague Centre For Strategic Studies, 2017.

DI CIOMMO F., *Evoluzione Tecnologica e Regole Di Responsabilità Civile*, Edizione scientifiche italiane, Napoli, 2003.

D'IGNAZIO G., *Il Sistema Costituzionale Degli Stati Uniti d'America*, Wolters Kluwer, 2020.

DREYFUS H., *What computers can't do: The limits of artificial intelligence*, Harper &Row, 1972.

FASCIONE L., *Storia Del Diritto Privato Romano*, Giappichelli, Torino, 2012.

FEIGENBAUM E.A.- FELDMAN J., *Computers and thought.* McGraw-Hill,1963.

FERRARI V., *Nuovi profili di diritto delle assicurazioni. Il fatto assicurativo*, Giuffrè, 2003.

FLOREANO D. – MATTIUSSI C., *Manuale Sulle Reti Neurali*, Il Mulino Bologna, 2002.

FRANCESCHETTI P., *La responsabilità civile*, Maggioli Editore, 2009.

FRANZONI M., *Fatti Illeciti: Art. 2043-2059*, Zanichelli, Roma, 2020.

FRIGIDA F., *Responsabilità del sorvegliante dell'incapace, dei genitori e tutori, dei padroni e Committenti*, in *La responsabilità civile*, Ed. Pasquale Fava, 1785-1816, Giuffrè Francis Lefebvre, 2018

GALGANO F., *Trattato di diritto civile*, Cedam, 2014.

GALLANT S.I., *Neural Network Learning and Expert Systems*, MITPress, Cambridge, MA, 1995.

GARFINKEL S., SPAFFORD G., *Web Security & Commerce*, Sebastopol, O'Reilly, 1997.

GIAUME, A. – GATTI S., *#Ai Expert: Architetti Del Futuro*, F.Angeli, Milano, 2019.

HÉNIN, S., *AI: Intelligenza Artificiale Tra Incubo e Sogno*, Hoepli, 2019.

HSU, FENG-HSIUNG. *Behind Deep Blue: Building the Computer That Defeated the World Chess Champion*. Princeton University Press, 2002.

IRTI N., *Le categorie giuridiche della globalizzazione*, in *Riv. dir. civ.*, 2002.

IRTI N., *Norma e luoghi. Problemi di geo-diritto*, Roma-Bari, 2006.

JACKSON P.C., *Introduction to artificial intelligence*, Dover Publications, 2019.

JACKSON P.C., *Toward human-level artificial intelligence: Representation and computation of meaning in natural language*, Dover Publications, 2019.

KAPLAN J., *Intelligenza artificiale: guida al futuro prossim*o. Luiss University Press, Roma, 2017.

LAGHI P., *Cyberspazio e sussidiarietà*, Napoli, ESI, 2015.

LAURA L., *Breve e universale storia degli algoritmi*, Luiss University Press, Roma, 2019.

LIDDY E.D., *Natural Language Processing*, in *Encyclopedia of Library and Information Science*, Eds. John D. McDonald and Michael Levine-Clark, 2001.

LOHSSE S. - REINER SCHULZE R. - STAUDENMAYER D., *Liability for Artificial Intelligence and the Internet of Things: Münster Colloquia on EU Law and the Digital Economy IV [Internet]*, Nomos, 2019.

MANTOVANI M., *La responsabilità dei genitori, dei tutori, dei precettori e dei maestri d'arte*, in *Commentario del Codice Civile dei fatti illeciti artt. 2044-2059*, Ed. Ugo Carnevali, 88-150, UTET giuridica, 2011.

MARMO R., *Algoritmi per l'intelligenza artificiale: Progettazione dell'algoritmo -Dati e Machine Learning - Neural Network - Deep Learning*, Hoepli, 2020.

MASI A., *Lezioni Di Istituzioni Di Diritto Romano*, Cedam, Padova, 2012.

MINSKY M. – PAPERT S.A., *Perceptrons: An Introduction to ComputationalGeometry,* MIT Press, Cambridge, MA, 1969.

MITCHELL T.M., *Machine Learning*, McGraw-Hill, New York, 1997.

NEGNEVITSKY M., *Artificial Intelligence: a Guide to Intelligent Systems*, Pearson Education, 2005.

NILSSON N.J., *Intelligenza Artificiale*, Apogeo,Milano, 2002.

NILSSON, N.J., *The Quest for Artificial Intelligence*, Cambridge University Press, 2009.

PAGALLO U., *The laws of robots: crimes, contracts, and torts*, Springer Science & Business Media, 2013.

PIREDDU M., *Algoritmi: Il Software Culturale Che Regge Le Nostre Vite*, L. Sossella, Roma, 2017.

POSNER R., *Economic Analysis of Law*, edit. Wolter Kluwer Law & Business, 2014.

RUFFOLO U. - ALPA G. - BARBERA A., *Intelligenza Artificiale: Il Diritto, i Diritti, l'Etica*, Giuffrè Francis Lefebvre, Milano, 2020.

RUFFOLO U., *Intelligenza Artificiale e Responsabilità: Responsabilità Da Algoritmo?;*

A.I. e Automobili Self-Driving, Automazione Produttiva, Robotizzazione Medico-Farmaceutica ; A.I. e Attività Contrattuali ; Le Tendenze e Discipline Unionali, Giuffrè Francis Lefebvre, 2017.

RUMELHART, D. E.- MACCLELLAND L., *Parallel Distributed Processing*, MIT Press, Cambridge, MA, 1986.

RUSSELL S.J. – NORVIG P., *Intelligenza Artificiale. Un approccio Moderno. Volume 1*, Pearson Prentice Hall, 2005.

SANTOSUOSSO A. – TOMASI M., *Diritto, Scienza, Nuove Tecnologie*, Wolters Kluwer, Milano, 2021.

SHORTLIFFE E., *Computer-based medical consultations: MYCIN. Vol. 2*, Elsevier,2012.

SISINI F., *Introduzione alle reti neurali con esempi in linguaggio C*, 2020.

SMOLENSKY P., *Il connessionismo tra simboli e neuroni* (traduzione di Frixione M.)., Marietti/Cambridge University Press, 1992.

SOMALVICO M., *Intelligenza artificiale*, Scienza & vita nuova, 1987.

STANZIONE P., *Responsabilità Extracontrattuale. Volume Secondo*, CEDAM, 2012.

STONE W., *The history of robotics*, in *Robotics and Automation Handbook*, ed. Thomas R. Kurfess, 1-12. CRC Press, 2005.

SUTTON R.S. – Barto A., *Reinforcement Learning: an Introduction*, The MIT Press, Cambridge, MA, 2018.

TADDEI ELMI G., *I diritti dell'intelligenza artificiale tra soggettività e valore. Ius condendum o fantadiritto?,* in *Il meritevole di tutela*, Ed. Luigi Lombardi Vallauri, Giuffrè Editore, 1990.

TEUBNER G., *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi* (traduzione di Femia P.), Edizioni Scientifiche Italiane, 2019.

TORRENTE A. – SCHLESINGER P., *Manuale Di Diritto Privato*, Giuffrè Francis Lefebvre, 2015.

TRAISCI F.P. - FIORE FONTANAROSA, *I diritti degli animali: da oggetti di consumo agroalimentare a soggetti giuridici con diritti propri*, in *Cibo e Diritto*, Ed. Lucia Scaffardi and Vincenzo Zeno Zencovich, 853-875, RomaTRE Press, 2020.

TRIMARCHI P., *Rischio e Responsabilità Oggettiva*, Giuffrè Editore, Milano, 1961.

WARWICK K. - BARATTIERI DI SAN PIETRO C., *Intelligenza Artificiale: Le Basi*, D. Flaccovio, Palermo, 2015.

WILSON, H. J. - DAUGHERTY P.R., *Human + Machine Reimagining Work in theAge of AI*, Harvard Business Review Press, Boston, MA, 2018.

ZATTI P. – COLUSSI V. – FUSARO A., *Lineamenti Di Diritto Privato*, Wolters Kluwer, 2020.

LAWS AND REGULATIONS

Costituzione italiana

Art. 13, 16, 42, 48, 53.

Codice Civile

Art. 1, 2, 172, 2043, 2046, 2047, 2048, 2049, 2050, 2051, 2740.

Decreto del Presidente della Repubblica 24 maggio 1998, n. 224.

Decreto Legislativo 30 giugno 2003, n. 196 (*Codice Privacy*).

Decreto Legislativo 6 settembre 2005, n. 206 (*Codice del Consumo*).

Art. 103, 115, 116, 117, 118

Regolamento UE 2016/679 (GDPR).

Direttiva 83/189/CEE del Consiglio del 28 marzo 1983 relativa alla "*procedura d'informazione nel settore delle norme e delle regolamentazioni tecniche*."

Direttiva 85/374/CEE del Consiglio del 25 luglio 1985 relativa al *ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati Membri in materia di responsabilità per danno da prodotti difettosi*.

Direttiva 2001/95/CE del Parlamento europeo e del Consiglio, del 3 dicembre 2001, relativa "*alla sicurezza generale dei prodotti.*"

Direttiva 2006/42/CE del Parlamento Europeo e del Consiglio del 17 maggio 2006 relativa "*alle macchine e che modifica la direttiva 95/16/CE*"

Risoluzione del Parlamento Europeo del 12 febbraio 2019, "*Una politica industriale europea globale in materia di robotica e intelligenza artificiale.*"

Risoluzione del Parlamento Europeo del 20 ottobre 2020, "*Regime di responsabilità civile per l'intelligenza artificiale".*

Proposal for a Regulation of The European Parliament and of The Council del 21 aprile 2021, "*Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts."*

COURT RULINGS

Cass. Civ., sez. Un., sentenza n. 19663/2014

Cass. Civ. sez. Un., sentenza n. 5422/2021

Cass. Civ., sez. I, sentenza n. 1931/2017

Cass. Civ. sez. II, sentenza n. 22728/2018

Cass. Civ., sez. III, sentenza n. 6148/1979

Cass. Civ., sez. III, sentenza n 11163/1990

Cass. Civ., sez. III, sentenza n. 5485/1997

Cass. Civ. sez. III, sentenza n. 1859/2000

Cass. Civ., sez. III, sentenza n. 6756/2001

Cass. Civ., sez. III, sentenza n. 7387/2001

Cass. Civ., sez. III, sentenza n. 8470/2001

Cass. Civ., sez. III, sentenza n. 11245/2003

Cass. Civ. sez. III, sentenza n. 8457/2004

Cass. Civ., sez. III, sentenza n. 12965/2005

Cass. Civ., sez. III, sentenza n. 15383/2006

Cass. Civ., sez. III, sentenza n. 4279/2008

Cass. Civ., sez III, sentenza n. 6632/2008

Cass. Civ., sez. III, sentenza n. 10741/2009,

Cass. Civ., sez. III, sentenza n. 2422/2014

Cass. Civ., sez. III, sentenza n. 17091/2014

Cass. Civ., sez. III, sentenza n.19657/2014

Cass. Civ., sez. III, sentenza n. 24475/2014

Cass. Civ., sez. III, sentenza n. 16052/2015

Cass. Civ., sez. III, sentenza n. 25223/2015

Cass. Civ., sez. III, sentenza n. 3258/2016

Cass. Civ., sez. III, sentenza n. 16661/2017

Cass. Civ., sez. III, sentenza n. 18865/2017

Cass. Civ., sez. III, sentenza n. 22058/2017

Cass. Civ., sez. III, sentenza n. 22839/2017

Cass. Civ., sez. III, sentenza n. 19180/2018

Cass. Civ., sez. III, sentenza n.29828/2018

Cass. Civ., sez. III, sentenza n. 5825/2019

Cass. Civ., sez. III, sentenza 8449/2019

Cass. Civ., sez. III, sentenza n. 13966/2019

Cass. Civ., sez. III, sentenza n. 25028/2019

Cass. Civ. sez. III, sentenza n. 4590/2020

Cass. Civ., sez. III, sentenza n. 8478/2020

Cass. Civ., sez. III, sentenza n. 9693/2020

Cass. Civ., sez. III, sentenza n. 25018/2020

Cass. Civ., sez. VI, sentenza n. 12410/2020

Cass. Pen., sez. IV, sentenza n. 38154/2009