# Introduction.

The aim of this paper is approaching the theme of Artificial Intelligence and the regulations involved when referring to non-contractual liability, specifically the damages that a new technological system may cause and both the rights of harmed individuals and the ones of any other party involved in it, following the whole chain from the creation of the product to its placing on the market and its usage.

In the first Chapter we will begin our coverage of the field by getting deep in the genealogy of it, how AI was firstly mentioned and idealised, through the scientific innovation that firstly brought it to life, arriving at the definitions of the main concepts evolved around it, or better, before it: the meaning of the word *"algorithm",* what is behind the study of *"machine learning",* understanding the gear, the mechanism of the Artificial Neural Networks and finally addressing the most efficient method for an AI system to become *"smart"* which is *"deep learning".*

In the second Chapter the theme of the non-contractual liability will come in place, starting with the fundaments of our civil law and the many principles around it, from the definition of Legal Subjectivity to the same concept of liability, reaching article 2043 which is the milestone for the compensational damage in our Country. Then we will continue by dividing our study about AI's liability into two parts, one where we see it as a sentient being, with all the complications around it and the vicarious articles that our civil code would analogically use to give accountability to the new system. And the second one, in which we will look at AI as a mere product without human's similarities

but addressing the problems regarding the necessary changes that adding something so different from "traditional" products would translate into. Specifically referring to the liability from the exercise of a dangerous activity, which is how some experts consider some high-risk technology, to the damage caused by inanimate things kept in custody and finding the person who would be liable for them; finally peaking with the discipline of the defective product, shaped by the EU, which is what will lead us directly into the international sovereignty in the field and so, the third and final Chapter.

In the last Chapter we will look at how Europe has decided to approach the theme, how it wanted, and still wants, to be at the forefront of such futuristic field, and take landmarking decisions that will shape the Union's future.

Specifically, we will begin by the Strategy proposed back in 2018 and what happened after the various Communications and the ambitious goals proposed by the European Commission. We will take a deeper look into the Resolution of the European Parliament of the 20th of October 2020, comparing it with the Artificial Intelligence Act proposed by the Commission in 2021 and how they do not necessarily complement each other, but better, work against one another. We will then try to understand what happened during the debates amongst doctrine, coming from all the State Members, the Governments, and the European and National Authorities, who were not too content with the outcome of the proposal for the new regulation and then, ending the research with the recent proposals made by the European Commission, trying to address the argument surged, of one new AI liability Directive and the update to the Defective Product's Directive of 1985.

One principle that has been significant in the last years and we will try to keep in mind during this study will be the fact that innovation and technological evolution is all around us, and as many authors during our study will refer to, it is hopeful to have balanced regulations on the field, now when we are still in the early stages, than finding ourselves with a critical void in the future.

# Chapter 1. Artificial Intelligence.

We live in a world resembling a lot Lewis Carrol's Wonderland described back in 1865. He wrote about "Symbolic Logic", which is analogous to "Propositional" or "Predicate Logic" used nowadays in AI. We often have no idea how everything started and how our whole environment could be like Orwell's "Big Brother" or even more so the one of Zemeckis' "Back to the future". Only recently, an actual timeline of 'how Artificial Intelligence was born' has been created, and we can understand and see the progress made over more than seven decades.

Every day we hear about big data, machine learning, deep learning, algorithms, human-like robots, self-driving cars, extremely risky robotic surgeries and even the possibility of having judgments decided by these automatons.[1] There are many, if not infinite ethical questions about what is right and what is wrong. Many experts say we will spend the next thirty years discussing regulations that will need implementation by countries and organisations worldwide. This way, we will decide if this is the kind of future we want and the limits to be imposed on the next generations.

In this first chapter, we will begin by analysing the history of Artificial Intelligence, the roller coaster that it has been: between peaks and falls; the meaning of it; its specifics; how it is always all around us even when we do not realise it and the prospects of the future on a technical and more mainstream level.

[1] U. RUFFOLO *"Per i Fondamenti Di Un Diritto Della Robotica Self-Learning; Dalla Machinery Produttiva All'auto Driverless: Verso Una 'Responsabilità Da Algoritmo'* in *'Intelligenza Artificiale e Responsabilità',* 2018, Torino.

## 1.1. Genealogy.

It is not easy to detect the beginnings of AI as it is hard with any phenomenon that was not discovered in one day but was the work and the study of many experts: mathematicians, scientists, and engineers; even famous literati such as the previously said Lewis Carrol who was also a mathematician and his, at the time, wild ideas have inspired many people after him who wandered how to transform his fantasy into reality. One specifically has been said to have settled the roots of AI, and for that, we need to go back to the ’40s and the short story *Runaround* by Isaac Asimov*,* which told us about a robot and the theory behind its creation: the imaginary “*Three Laws of Robotics:*

1. *A robot may not injure another human being or, through inaction, allow a human being to come to harm;*
2. *A robot must obey the orders given to it by human beings except where such orders would conflict with the First Law;*
3. *A robot must protect its own existence as long as such protection does not conflict with the First and Second Laws”.*

On the other side of the world, in 1950, the renowned mathematician Alan Turing (remembered to have allowed the British Government to break the Enigma code used by the German army during the Second World War) developed a machine called “*The Bombe”* that is considered the first working computer. This creation led Turing to fantasise about the intelligence of these machines. Later, he published an article, “*Computing Machinery and Intelligence*”, outlining how to make these smart machines. Still, most famous is his *“Touring Test”,* a benchmark to understand the intelligence of an artificial system: “*if an evaluator is interacting with another human and a machine,*

*and the first is unable to distinguish the machine from the human, then the machine is said to be intelligent".*[2]

The word was officially conceived six years later, in 1956, during the eight-week-long *Dartmouth Summer Research Project on Artificial Intelligence,* also known as *"DSRPAI", at* Dartmouth College in New Hampshire. This seminar is known to have put together those who would be considered later as the founding fathers of AI, specifically Marvin Minsky and John McCarthy, who hosted the event. The goal was to assemble various researchers from different areas of expertise to create a spark for the innovative idea of machines that would emulate human intelligence. It was an epic moment in the history of AI, an event that gave us its first definition: "*a science and a set of computational technologies that are inspired by – but typically operate quite differently from – the ways people use their nervous systems and bodies to sense, learn, reason and take action"*[3].

An example of one of the first computer programs to follow this definition was the famous ELIZA, created between 1964 and 1966 by Joseph Weizenbaum at MIT. It passed the Turing Test and had the objective of simulating a conversation with a human, followed by the General Problem Solver developed by Herbert Simon, Cliff Shaw and Allen Newell, which was able to solve on its own some simple problems.

Three years later, despite the success of the following years and the incrementing of the funding given by different associations and the American Government, politics got in the middle of the argument and started to highly criticise all the money invested in AI research. The lack of progress in the field of AI lies in the way that those systems such as

[2] M. HAENLEIN *"A Brief History of Artificial Intelligence On the Past Present and Future of Artificial Intelligence."* www.researchgate.net

[3] G. F. ITALIANO*, "Intelligenza Artificiale: Passato, Presente, Futuro"* in Pizzetti F. (a Cura Di) *"Intelligenza Artificiale, Protezione Dei Dati Personali e Regolazione"* 2018, Milano.

ELIZA and IBM's *"Deep Blue"* chess-playing program were working, assuming to formalise "human intelligence" in an approach such as "if-then" statements; clearly showing the limit of the programs as well highlighted by James Lighthill who said that they would have only reached the level of experienced amateurs in games and that common sense reasoning would always be beyond their abilities. Although his statement was proven half wrong many years later when IBM's program was able to beat the Chess world champion, Gary Kasparov, what these machines needed was the ability to replicate the process of neurons in the human brain, a theory that in the 40s had been put out by the Canadian psychologist Donald Hebb already. Marvin Minsky showed that computers did not have sufficient processing power to handle all the work of creating an artificial neural network. In addition, various issues were added to the argument by different experts: Richard Karp, in 1972, showed that only in extended time could research solve many problems and that it would have needed an exponential amount of time for a computer[4]; researchers soon understood that to do more these machines would need applications of live vision or natural language, at least the knowledge that a kid would have of its surroundings[5]; *the Moravec's paradox* showed that one thing was to resolve mathematics problems and another would have been recognising a face or moving into a room without hitting obstacles[6]; John McCarthy in particular underlined that they were not going to be able to represent or solve standard deductions involving planning or reasoning without making changes to the structure of the logic itself[7]. This time between the 60s and 70s is considered the "winter" of AI because of the many falls that the research

[4] S.J. RUSSEL. P. NORVIG, "*Artificial Intelligence: A Modern Approach*"(Upper Saddle River, New Jearsey: Prentice Hall, 2003).
[5] 'P. McCORDUCK, *"Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence"* Pp. 300 & 421, 2004; D. CREVIER *"AI: The Tumultuos History of the Search for Artificial Intelligence"* 1993, Pp. 113–114; H. MORAVEC *"Mind Children: the future of Robot and Human Intelligence"* 1988, p. 13.
[6] P. McCORDUCK 2004, p. 456, H. MORAVEC 1988, Pp. 15–16'.
[7] 'J. McCARTHY, P.J. HAYES *"Some Philosophical Problems From The Standpoint of Artificial Intelligence"* Edinburgh University Press. pp. 463--502, 1969,.

was going through, and soon it became pretty clear that the goals projected years before were far from becoming a reality:

1. The frustration by the different organisations such as DARPA (research and development agency) and NRC (collective scientific national academy) in the United States and by the British Government, the two countries most focused on research during the 70s which were funding many projects but with no actual improvement nor discoveries. It is reported that after spending 20 million dollars only from the NRC, by 1974, a way to fund AI projects was hard to find.[8] Nonetheless, DARPA was under constant pressure by the Government to fund direct research; thence, the endowment turned towards projects with clear objectives, such as autonomous tanks and battle management systems.[9];
2. Moreover, several philosophers criticised the goals of AI researchers. Hubert Dreyfus argued that the human reasoning they wanted to give machines was made mainly by embodied, automatic and unconscious "know-how". [10] All critiques that were not taken seriously by AI researchers for being off the point; one of them, though, the author of ELIZA, Weizenbaum, was particularly stroked by ethical questions when the psychiatrist Kenneth Colby suggested that a *"computer program [...] can conduct a psychotherapeutic dialogue"* based on Eliza.[11] He

[8] 'McCORDUCK 2004, Pp. 280–281, CREVIER 1993, p. 110, RUSSELL & NORVIG 2003, p. 21 and NRC 1999 under *"Success in Speech Recognition"*.'

[9] NRC 1999 under *'Shift to Applied Research Increases Investment.'* While the Autonomous Tank Was a Failure, the Battle Management System (Called 'DART') Proved to Be Enormously Successful, Saving Billions in the First Gulf War, Repaying the Investment and Justifying the DARPA's Pragmatic Policy, at Least as Far as DARPA Was Concerned.

[10] *'Know-How'* Is Dreyfus' Term. *(*Dreyfus Makes a Distinction between *'Knowing How'* and *'Knowing That', a* Modern Version of Heidegger's Distinction of Ready-to-Hand and Present-at-Hand*.) (Dreyfus & Dreyfus 1986)*; Dreyfus' *"Critique of Artificial Intelligence"*

[11] COLBY, WYATT & GILBERT *"A Computer Method of Psychoterapy"* 1966, p. 148. Weizenbaum Referred to This Text in Weizenbaum 1976, Pp. 5, 6. Colby and His Colleagues Later Also Developed Chatterbot-like *'Computer Simulations of Paranoid Processes (PARRY)'* to *'Make Intelligible Paranoid Processes in Explicit Symbol Processing Terms.'* (Colby 1974, p. 6).

was so disturbed by this that he published *'Computer Power and Human Reason'* asserting the potential devaluing of human life that could come from the misuses of artificial intelligence. [12];

3. Previously in 1958, Frank Rosenblatt introduced the *"perceptron",* a form of neural network that led to 'Connectionism', a range of techniques and algorithms using artificial neural networks.[13] Minsky and Seymour later strongly critiqued it. Lastly, Papert's publication of *"Perceptrons"* in 1969 publicly destroyed his theory (being only reborn more than ten years later and eventually becoming vital and valuable to artificial intelligence today). It was clear evidence of a battle even amongst the same researchers. [14]
4. Even the logic and symbolic approach previously introduced by John McCarthy in his *"Advice Taker proposal"* and later brought back to life by Robert Kowalski at the University of Edinburgh in the 70s was strongly criticised by Dreyfus and famous psychologists such as Peter Wason and Eleanor Rosch. This technique implemented deduction and solved problems through the resolution and unification algorithm that led to the creation of a successful logic programming language called *'Prolog'*.[15] The critics once again noted that human beings rarely used logic to solve problems, but instinct, to which McCarthy replied that the only thing needed was for machines to solve and not to think as people do.[16] His approach will be referred to as *"neat",* as opposed to the *"scruffy"* one of other

[12] Weizenbaum's Critique of AI: McCORDUCK 2004, Pp. 356–373, CREVIER 1993, Pp. 132–144, RUSSELL & NORVIG 2003, p. 961 and See WEIZENBAUM *"Computer Power and Human Reason: From Judgment to Calculation"* 1976.
[13] P. SMOLENSKY (1999). *'Grammar-Based Connectionist Approaches to Language'*. Cognitive Science. 23 (4): 589–613. Doi:10.1207/S15516709cog2304_9.
[14] McCORDUCK 2004, Pp. 104–107, CREVIER 1993, Pp. 102–105, RUSSELL & NORVIG 2003, p. 22.
[15] CREVIER 1993, Pp. 193–196.
[16] An Early Example of McCarthy's Position Was in the Journal Science Where He Said *'This Is AI, so We Don't Care If It's Psychologically Real'* (Kolata 2012), and He Recently Reiterated His Position at the AI@50 Conference Where He Said *'Artificial Intelligence Is Not, by Definition, Simulation of Human Intelligence'* (Maker 2006).

colleagues of him, who were highly against him and were trying to see better into the *"story understanding"* and *"object recognition"* that a machine needed to think like a human being.[17] Minsky, in particular, was against McCarthy (who was the leader of the "scruffy" movement) and noted that he and his fellow researchers were using the same instrument: capturing all of our ordinary common sense assumptions about something, referring to these as *"frames".*

In the early 80s, things drastically changed positively due to: (i) the rise of an AI program called *"expert system"* that was adopted around the world by many corporations; (ii) the extreme funding of the Japanese Government towards AI with its *"fifth generation computers"* propaganda; (iii) the comeback of Connectionism.

(i) This program was able to answer and solve problems of a specific domain of cognizance, in this way, avoiding the common sense knowledge problem and using the awareness derived from a certain amount of experts in that area. An example was *"Dendral",* a project of AI that in 1972 was able to diagnose infectious blood diseases.[18] These programs proved themselves very useful, something AI had still not achieved. The success was incredible: an expert system named *'XCON'* in the United States was successfully implemented for the *"Digital Equipment Corporation",* and by 1986 it was saving the Government 40 million dollars annually. [19] Historically speaking, we can refer to this moment as *"the knowledge revolution",* a new direction taken by AI

[17] 'Neat vs. Scruffy: McCORDUCK 2004, Pp. 421–424 (Who Picks up the State of the Debate in 1984). CREVIER1993, Pp. 168 (Who Documents Schank's Original Use of the Term). Another Aspect of the Conflict Was Called *"the Procedural/Declarative Distinction"* but Did Not Prove to Be Influential in Later AI Research.'

[18] 'McCORDUCK2004, Pp. 327–335 (Dendral), CREVIER 1993, Pp. 148–159, NEWQUIST *"The Brain Makers"* 1994, p. 271, RUSSELL & NORVIG 2003, Pp. 22–23'.

[19] CREVIER 1993, p. 198'.

researchers in which *"intelligent behaviour depended very much on dealing with knowledge, sometimes quite detailed knowledge, of a domain where a given task lay"* [20], becoming the focal point for all researchers. An important example of this focus in the 80s was the birth of *'Cyc'*, an endeavour to contrast the previous common sense knowledge problem. Douglas Lenat, leading the project, said that the method to follow was for machines to understand human concepts and to do so, they had to teach them, one notion at a time, hand in hand, like with a baby; it was clear very soon though that this was going to take decades. [21]

(ii) The second significant event was the massive funding that the Japanese Ministry of International Trade and Industry decided to reserve for the "Fifth generation computer": a ten-year enterprise to build some sort of *"Supercomputer"* with the target of fabricating multiprocessors able to perform conversations, translations and reasoning like human beings [22], choosing 'Prolog' as the essential language for the project.[23] At the same time, the rest of the world responded to this advanced research with their own investing: the UK funded the *'Alvey'* project with 350 million pounds. A pool of American corporations formed the *'Microelectronics and Computer Technology Corporation*, aka MCC'[24] and DARPA additionally founded the *"Strategic Computing Initiative",* tripling its financing to AI projects.[25]

[20] 'McCORDUCK 2004, Pp. 421'.
[21] McCORDUCK 2004, p. 489, CREVIER 1993, Pp. 239–243, NEWQUIST 1994, Pp. 431–455, RUSSELL & NORVIG 2003, p. 363−365 and LENAT & GUHA *" Building Large Knowledge-based Systems: Representation and Inference in the Cyc Project"*1989.
[22] McCORDUCK 2004, Pp. 436–441, NEWQUIST 1994, Pp. 231–240, CREVIER 1993, Pp. 211, RUSSELL & NORVIG 2003, p. 24 and See Also FEIGENBAUM & McCORDUCK *"The fifth generation: Artificial Intelligence and Japan's Computer Challenge to the World" 1983*.
[23] 'CREVIER 1993, Pp. 195'.
[24] CREVIER 1993, Pp. 240.'; 'RUSSELL & NORVIG 2003, p. 25'.
[25] 'McCORDUCK 2004, Pp. 426–432, NRC 1999 under *"Shift to Applied Research Increases Investment"*.