FONDAZIONE LEONARDO
Civiltà delle Macchine

CIVILTÀ DELLE MACCHINE

CENTRO STUDI AMERICANI

Winning the Artificial Intelligence era

QUANTUM DIPLOMACY AND THE POWER OF AUTOMATION

INDEX

PREFACE 6

INTRODUCTION

- 1. THE ERA OF DEEP TECHNOLOGIES 7
- 2. IMPACT ON GEOPOLITICS AND INTERNATIONAL RELATIONS 9

CHAPTER I. CAN WE QUANTIZE INTERNATIONAL RELATIONS?

by Enrico Prati

1. DECODING SOCIAL SYSTEMS 14

2. INTERNATIONAL RELATIONS: WHEN CHANGING THE ORDER OF THE FACTORS, THEN THE RESULT DOES CHANGE 17

- 2.1 QUANTUM THEORY, SCIENCE AND TECHNOLOGY 18
- 2.2 THE APPLICATION OF THE CATEGORIES OF THOUGHT OF QUANTUM THEORY TO INTERNATIONAL RELATIONS 19

3. CRITICAL CONSIDERATIONS ON THE QUALITATIVE QUANTIZAZION OF INTERNATIONAL RELATIONS 22

CHAPTER II. QUANTUM TECHNOLOGIES AND INTERNATIONAL RELATIONS

by Enrico Prati

1. TOWARDS THE INTEGRATION OF QUANTUM TECHNOLOGIES 27

2. GEOPOLITICS AND QUANTUM TECHNOLOGIES 30

3. QUANTUM COMPUTERS: HYPE OR TECHNOLOGY PLATFORM FOR SOCIAL SCIENCE? 34

4. A CASE STUDY: THE EVOLUTION OF GLOBAL TERRORIST NETWORKS 37

CHAPTER III. THE USE OF ARTIFICIAL INTELLIGENCE IN WEAPONS SYSTEMS: INTERNATIONAL LEGAL FRAMEWORK

by Andrea Gilli and Lucrezia Scaglioli

1. ARTIFICIAL INTELLIGENCE: DEVELOPMENT AND SCEPTICISM **41**
2. TECHNOLOGICAL REVOLUTION **42**
3. MILITARY IMPLICATIONS **44**
4. ETHICAL AND LEGAL QUESTIONS **49**

CHAPTER IV. NEW WARFARE: POTENTIAL RISKS AND MITIGATIONS

by Enrico Savio and Enrico Comin

1. DISRUPTIVE TECHNOLOGIES IN DEFENCE **55**
2. ARTIFICIAL INTELIGENCE (AI) **56**
3. QUANTUM TECHNOLOGIES **59**
4. LETHAL AUTONOMOUS WEAPON (LAWS) **61**
5. HYPERSONIC **63**
6. DIRECT ENERGY WEAPONS (DEW) **65**
7. SELF-DRIVING SYSTEM **66**
8. THE IMPLICATIONS AND IMPACT OF THESE TECHNOLOGIES ON CURRENT TACTICAL AND STRATEGIC PARADIGMS: THE TRANSITION TO HYPERWAR **68**
9. DETERRENCE IN THE FUTURE HYPERWAR CONTEXT **72**
10. DISRUPTIVE TECHNOLOGIES, DEFENCE AND SECURITY: ETHICAL AND MORAL ISSUES **76**

CHAPTER V. LEGISLATION ON ARTIFICAL INTELLIGENCE OUTSIDE THE EUROPEAN UNION: A COMPARATIVE ANALISYS

by Antonio Malaschini

1. THE GLOBAL REGULATORY FRAMEWORK **83**

2. CHINA **84**

3. UNITED STATES **86**

- 3.1 REGULATORY INITIATIVES **87**
- 3.2 THE CONCLUSIONS OF THE NATIONAL SECURITY COMMISSION ON AI **89**

4. UNITED KINGDOM **91**

5. RUSSIA **93**

6. A FIRST COMPARISON **94**

CHAPTER VI. LEGISLATION IN THE EUROPEAN REFERENCE FRAMEWORK AND THE ITALIAN CASE

by Antonio Malaschini

1.INTRODUCTION **98**

2. THE EUROPEAN COMMISSION'S DRAFT REGULATION **99**

- 2.1 THE RISK CRITERION **100**
- 2.2 GOVERNANCE AND SANCTIONS **101**
- 2.3 THE INITIAL CRITICAL OBSERVATIONS **102**

3. ITALY **104**

4. CONCLUSIONS **106**

PREFACE

Artificial Intelligence, big data analysis and quantum computers will lead to an all-round, increasingly pervasive revolution. The availability of exponentially greater computing power will allow decision-makers to assess and act with an unprecedented context awareness and speed; a near future where the capabilities of machines may well exceed the human element, its control, its responsibility.

Resulting from the collaboration between Centro Studi Americani and Fondazione Leonardo-Civiltà delle Macchine with the support of the Italian Ministry of Foreign Affairs and International Cooperation and with the scientific contribution of Prof. Enrico Prati - this study aims to offer as objective an analysis as possible on the state of the art of new technologies and their applications in the new Defence and Security scenarios. It also aims to delve into and analyse the risks, main challenges and possible strategies to manage the developments already underway, within the framework of shared standards and values.

Like innovation, which - by its very nature - is an ongoing, complex, non-linear process, the definition of moral and ethical principles can only emerge from a constant dialogue between science, technology, as well as legal and political institutions. This is a process that puts the human at the centre as the ultimate decision-maker, who negotiates and sets the rules of the game, in every field of application, in times of peace and in conflict scenarios.

INTRODUCTION

by Enrico Prati

1. The era of deep technologies

Looking at the rapid evolution of electronic and information technologies that started in the 1970s, there is no doubt that starting from 2015-2016 we have been witnessing a new innovation momentum that is further accelerating the already rapid rate of change of technologies. We already find ourselves in the materialisation of many science fiction depictions of fifty or sixty years ago: we can videocall our loved ones, we are surrounded by an automated home environment, or we travel on driverless means of transport. This is not only our new normal, but we see that the change is only just beginning.

Among the drivers of acceleration, disciplines such as quantum mechanics and artificial intelligence stand out for their ever increasingly crucial role in various fields of application. Although they were founded respectively approximately one hundred years ago and seventy years ago, and we been exploiting them for decades, it is only in the last few years that we have witnessed their new and modern use. Google's quantum advantage computer announced in 2019 or deep learning applied to the recognition of retinal damage or tumours are just two examples. These innovations are generalist, as was the computer, and open up new application horizons potentially in all areas of knowledge. Due to both the high degree of specialisation required, and the speculative mathematical component involved as well as the advanced engineering of materials and techniques- not to mention the expectation that they may lead forward to new unforeseen horizons, e.g. the invention of the World Wide Web), technologies based on these disciplines are included in the sphere known as deep technologies - deep tech, borrowing an attribute of deep neural networks, which have recently opened up many new frontiers.

One fact that may come as a surprise is that quantum mechanics originated as a mathematical theory to describe very small systems, such as electrons and atoms, and its applications have been known to us for a long time. If its formulation and its relationship to physical reality had not been understood earlier, the realisation and exploitation, for example, of semiconductors and lasers in the 1960s would not have possible. Similarly, artificial intelligence, a mathematical theory applied to data and information, boasts a glorious tradition from its origins that has led to advanced algorithms such as expert systems or to

Deep Blue, the program that in 1997 beat Garry Kasparov, then the reigning world champion in the game of chess.

Investment in deep tech is primarily supported by governments, but also by the private sector of venture capitalists and corporations such as Google, Microsoft, IBM, Intel, Blackberry or Alibaba, to name a few. It has amplified the creation of new knowledge, generating the acceleration that we are currently witnessing, and which has given new impetus to technological innovation. In the United States, as well as in Canada, Israel, the United Kingdom and China, to name but a few countries that are particularly representative from the point of view of public innovation policies, the state has assumed the role of ecosystem creator, while investors and corporations are playing the role of selectors of the ideas, which are most likely to succeed from a market perspective, in sectors such as aerospace, health, security and so on.

On the public side, the creation of ecosystems includes strong support for the training of qualified human capital through the strengthening of universities, tax incentive policies for companies spread over at least ten years, the creation of technology parks, and the funding of basic research. On the private side, the financing of technology transfer to the market requires open innovation policies, availability of venture capital, and investment by corporations in research programmes in their R&D units of many hundreds of millions of euros.

These ingredients, combined together, have actually led to the development and exploitation of new hardware and new software that are making the fortunes of already established hardware manufacturers. To give one example, Nvidia from 2006 to 2021 has increased its value a hundredfold on the stock exchange. Then there are new start-ups that did not exist until five years ago and are now listed on NASDAQ, such as the quantum computer manufacturer IonQ whose value is now a few months away from its IPO of USD 2.8B, or the cybersecurity company employing artificial intelligence, Crowdstrike, founded in 2011 and worth USD 1 billion in 2017, up to its current value of more than USD 50 billion.

However, this is only the beginning. While ethical challenges on the regulatory side, are already arising – (e.g. the artificial intelligence-piloted military drones developed by some countries such as China and Turkey) these technologies - on the purely technical side - promise new scenarios. These range from the use of new materials as a physical substrate to support the abstract architecture of computation, to general artificial intelligence capable of emulating human intelligence and brain-machine interfaces such as those of Elon Musk's Neuralink - which raised a new round of funding of USD 205 million in 2021 from Amazon and Google. In the latter, artificial electronic grafts in the cortex of the human brain allow for a two-way exchange between the thought that takes place in the encephalon (biological neurons), and additional sense organs based on semiconductor and metal sensors, mediated by a system of artificial neurons trained with modern artificial intelligence methods. It is no coincidence that

DARPA has included BMI in the USD 2 billion 'AI Next' funding of the third generation of artificial intelligence, the one that is to go beyond today's deep learning.

Quantum computers and artificial intelligence are providing new computational power - to paraphrase Tom Conte of IEEE Rebooting computing, “new hardware for new types of software” - that will sustain economic growth and alter geopolitical balances for many years to come.

2. Impact on geopolitics and international relations

There is no doubt that this technological tsunami is altering the geopolitical balances, as proved bythe implications of the concentration in the East of the crisis in the availability of processors and integrated circuits (in particular with TSMC in Taiwan and Samsung in South Korea, which alone produce 70% of the world's semiconductors). This prompted US President Joe Biden to allocate USD 52 billion to boost domestic chip production capacity and the European Union to launch the EUR 45 billion Chips Act in February 2022. Analysing how these technologies may influence geopolitics and international relations, we can find two levels of impact. The first one is their direct use, through the application of technology as a tool to manage diplomatic relations. The second level concerns their indirect use, i.e. - in a proactive approach- as a driver to increase industrial and economic competitiveness, but also, in pragmatic terms, as an instrument of technological supremacy to exert greater pressure in diplomatic terms.

Focusing on direct employment, several scenarios can be listed as examples. Artificial intelligence can make it possible to determine the moves an artificial agent must make to strengthen its position in the representation of a diplomatic scenario. Its ability to handle large amounts of data can allow the quantitative analysis of an evolving scenario and predict its subsequent evolution in terms of the indicators considered. It can support human decision-making, ranging from the automation of routine tasks to support at the tactical and operational level, though presumably without reaching the autonomous strategic decision-making, which would remain the prerogative of humans[1].

Furthermore, by taking the mathematics underlying quantum mechanics to its extreme consequences, it is possible to formulate problems concerning diplomatic scenarios using concepts exclusive to quantum mechanics, such as that of quantum superposition, which, when transferred to a scenario, can be made to correspond to the fact that, until a decision has been made, two

[1] Bjola, Corneliu. "Diplomacy in the Age of Artificial intelligence." EDA Working paper, *Retreived from http://www.realinstitutoelcano.org/wps/portal/rielcano_en/contenido* (2019).

opposing alternatives coexist simultaneously, while only after the decision has been made will one of them be realised. Today, thanks to quantum processors, it is no longer a matter of the qualitative application of theory to the formulation of problems - which is a challenge in itself - but also potentially of the ability to solve them, once properly set up, by means of customised hardware ready to return an answer of interest. Representative fields of application are quantum decision theory or the simulation of community dynamics.[2] Particularly illustrative in this respect is the use of a quantum computer used to study the stability of network formation between terrorist factions in Iraq and Syria from 2006 to 2016. [3]

We are reminded by James Der Derian, a pioneer in the field, that the term quantum diplomacy originated from a conversation between the physicist Sidney Drell and President Reagan's Secretary of State George Shultz. The former observed that when in physics you observe something, the object of the observation changes. The latter replied that you only need to put a camera to observe something to immediately realise that the observation itself will be influenced. Regardless of the fact that the concept of quantum diplomacy is still being defined and requires a path of rapprochement between such distant topics as quantum physics and international relations, a potentially even long path, we are in fact witnessing a transition from an era in which diplomats engaged in regulation of technology (such as the nuclear disarmament negotiations), to one in which they will also engage in positive terms with new technology (around green diplomacy to meet the climate challenge together) as well as regulation on new weapons that will also be based on it. In this scenario, one can agree with those who argue, like Sem Fabrizi, that diplomacy must remain human-led even though in the future it is likely to be assisted by new disruptive technologies, including quantum technologies, provided that the diplomat remains at the centre of the analysis.

As far as indirect use is concerned, we can still give some concrete examples. There is an artificial intelligence competition between the US and China[4], which in recent years has accelerated its pursuit and in some cases even surpassed the US in terms of, for example, patent production or funding in certain areas. In response, former US President Donald Trump signed Executive Order 13859 in 2019 to maintain American leadership, stating that, “Continued American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in

---

[2] Lucas, Robert F., et al. "Practical Adiabatic Quantum Computing: Implications for the Simulation Community." *the Proceedings of the Interservice/Industry Simulation, Training and Education Conference, Orlando, Florida*. 2013.

[3] Ambrosiano, John Joseph, Randy Mark Roberts, and Benjamin Hayden Sims. *Using the D-Wave 2X quantum computer to explore the formation of global terrorist networks*. No. LA-UR-17-23946. Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2017.

[4] Kļaviņš, Didzis. "Diplomacy and Artificial Intelligence in Global Political Competition." *Global Studies* (2021): 213.

a manner consistent with our nation's values, policies, and priorities". Even in the field of quantum computers, China is challenging the supremacy of the United States[56] for both a mere reason of prestige (in fact, to date there is no hardware powerful enough to solve real large-scale problems, even though the roadmaps foresee quantum advantage in the next few years) and for a contingent reason. Being faster than the opponent to calculate the scenario more quickly leads to a significant position advantage since it is like predicting the future.

We don't need a crystal ball to guess that the control of the geopolitical scenario will be the prerogative of the powers that will have best been able to benefit from the application of deep tech technologies and advanced methods of analysis, which will prove more powerful than those of their competitors (or opponents). Conversely, these technologies bring any benefit to society if they are not accompanied by a conscious maturing of the ethical sphere, the social component, transparency and respect for rights and privacy. This is an interdisciplinary challenge that requires the commitment of all stakeholders, and a synthesis that only politics, the art of due measure as Plato interpreted, can deliver.

[5] Han-Sen Zhong et al, Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light, Physical Review Letters (2021). DOI: 10.1103/PhysRevLett.127.180502

[6] Yulin Wu et al, Strong Quantum Computational Advantage Using a Superconducting Quantum Processor, Physical Review Letters (2021). DOI: 10.1103/PhysRevLett.127.180501

Chapter I

CAN WE QUANTIZE INTERNATIONAL RELATIONS?

by Enrico Prati

SUMMARY: We can apply the categories of thought inherent in quantum theory to the social sciences, including international relations. *Quantum social* science is the science that aims to investigate the problems of the social sciences, be they economics, finance, psychology, sociology, with the help of formal methods developed in quantum theory. The concepts of discontinuity (quantization), complementarity, uncertainty, and so on, are suitable for qualitatively describing social and economic phenomena including those inherent in international relations, such as the exchange of money, a referendum towards a secession, the national identity and the state of relations between countries. These qualitative analogies form a basis for even quantitative modelling, which allows predictive analysis as occurs in the case of *quantum decision theory*.

1. DECODING SOCIAL SYSTEM

For some years now, a reflection has been in the making in the diplomatic sphere around the goal of quantifying international relations, in the sense of applying the fundamental concepts of quantum theory to them. This combination of a branch of the social sciences and a fundamental theory of physics offers, on the one hand, stimulating insights for both disciplines and prolific interaction, but at the same time lends itself to possible abuse of language, misinterpretation and overestimation. As happened in the past between physics and philosophy in order to achieve appreciable results on a gnoseological, methodological and even predictive level, it is necessary for the representatives of the two disciplines, with different and complementary backgrounds, to take a step towards rapprochement by first constructing such a common language, lending the discussion the rigour of their arguments, but in the awareness of their own limitations when progressively moving away from their own field of competence. The first step consists in giving a shared meaning to the expression quantifying international relations by retracing the path already started in the recent past by some pioneers in the field, not without having first introduced the fundamental concepts - also in a historical perspective - brought by what is a candidate to be an evolved tool for analysing and predicting geopolitical scenarios, namely quantum theory.

In the early years after its emergence, consolidated in the late 1920s, quantum mechanics appeared more as a puzzle of mathematical formulations united by a general approach and applied to a series of experiments, rather than a unifying and all-encompassing theorisation. For this reason, among the contributions of founding fathers such as Heisemberg, Schroedinger, Bohr, Fermi, Dirac and

Pauli, that of John Von Neumann, a Hungarian physicist naturalised in the United States, stands out. In fact, Von Neumann was able to formulate the axioms of quantum mechanics that are still taught today in a single framework, such that they encompassed everything that had been formalised until then in a coherent manner[7]. Representative in this context, however, is the further fact that he was not only able to mathematically formulate the theory of quantum mechanics, but that we also owe to him the invention of game theory in the economic sphere[8], which was later taken up and continued by economist John Nash[9]. It is no coincidence that the - albeit exceptional - person who formalised quantum mechanics is at the same time the one who applied mathematics to systems in which rational choices, be they economic, political or social, have to be calculated. In fact, systems in which, for example, a reward and punishment mechanism is involved can be formulated quantitatively, and many analysis tools developed in physics to describe experimental systems also find direct application in social systems.

The fact that specifically quantum mechanics' own methods can be converted to better explain - and thus necessarily also more accurately predict - the phenomena described by the disciplines falling under the social sciences - thus an area even much broader than international relations - has been the subject of academic study since 1978. In that year, the mathematical physicist Asghar Qadir theorised that the decision-making process can best be described by a quantum mathematical approach[10] rather than by modelling the decision-maker mechanistically, i.e. subject to forces based on utility and disutility, as assumed for the rational economic man of neoclassical economics.

A compendium of such applications has been formulated by Haven, Khrennikov and Khrennikov[11], with an excursus that starts firstly from the very definition of quantum social science as that which "*...has as goal to investigate problems within the wide remit of the social sciences; be it economics, finance, psychology, sociology or other domains of inquiry; with the help of formal models and concepts used in quantum physics*". The book identifies four categories of quantitative problems that benefit from the use of the mathematical methods of quantum mechanics, namely asset pricing in finance, decision making, quantum game theory, and a category of new concepts in the social sciences, including for example the application of Heisemberg's uncertainty principle to social systems[12].

---

[7] Birkhoff, G., & Von Neumann, J. (1936). The logic of quantum mechanics. *Annals of mathematics*, 823-843.

[8] Von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior*. Princeton university press.

[9] Nash, J. (1951). Non-cooperative games. *Annals of mathematics*, 286-295.

[10] Qadir, A. (1978). Quantum economics. *Pakistan Economic and Social Review*, *16*(3/4), 117-126.

[11] Haven, E., Khrennikov, A., & Khrennikov, A. I. (2013). Quantum social science. Cambridge University Press.

[12] Baaquie B. (2005). Quantum finance. Cambridge University Press; Cambridge.

Athalie and Haven[13] pick up on and take this analysis further, identifying from the 21st century onwards a current of interdisciplinary scientific thought that aims to develop mathematical models employing the tools originally conceived to describe quantum reality in order to explain certain socio-economic processes and human behaviour. An essential component that is used in such applications is the probabilistic nature of quantum mechanics, which means that it does not produce certain results but estimates the probability of a given event occurring. Again, the fields of application range from decision making to finance, but extend into the analysis of socio-economic data by considering the same physical rather than purely mathematical aspects of quantum physics.

In the wake of this innovative approach to the social sciences are the pioneering contributions of Der Derian[14] and Hundt[15] respectively, who combine these new methodologies with the specific field of international relations. The hypothesis is that international relations and the geopolitical sphere in general lend themselves to a conceptualisation that can benefit from an application of the principles governing quantum mechanics. In the remainder of this chapter, we describe the conceptual affinities between quantum mechanics and international relations, the areas of application that have been reported in the literature, and some critical considerations and limitations on the validity of the application of this approach. Referring again to the search for the estimated probability of a given event, be it a military crisis, a collapse of the markets[16], or the default of a sovereign state, this is a central aspect of the preconditions for a decision in the international arena. This Chapter I specifically examines quantum diplomacy with a conceptual and qualitative slant, inherent in the application of the categories of thought derived from the peculiar aspects of quantum theory. Chapter II examines the quantitative application of quantum theory to the social sciences and in particular to the case of international relations.

Both chapters are developed in accordance with two conceptual cornerstones, the first of which - echoing an observation by Der Derian and Wendt themselves - that in the face of the technological wave that is manifesting itself through the new quantum technologies, it is better to anticipate than to follow the emergence of the phenomenon by preparing for the language and tools that will be appropriate to it.

---

[13] Athalye, V., & Haven, E. (2021). Socio-Economic Sciences: Beyond Quantum Math-like Formalisms. Quantum Reports, 3(4), 656-663

[14] Der Derian, J., & Wendt, A. (2020). 'Quantizing international relations': The case for quantum approaches to international theory and security practice. Security Dialogue, 51(5), 399-413.

[15] Wendt, A. (2015). Quantum mind and social science. Cambridge University Press.

[16] Ding, Y., Gonzalez-Conde, J., Lamata, L., Martín-Guerrero, J. D., Lizaso, E., Mugel, S., ... & Sanz, M. (2019). Towards prediction of financial crashes with a D-Wave quantum computer. *arXiv preprint arXiv:1904.05808*.

2. INTERNATIONAL RELATIONS: WHEN CHANGING THE ORDER OF THE FACTORS, THEN THE RESULT DOES CHANGE

In order to create a common language between practitioners of international relations and scholars of quantum sciences and technologies, it is important to make it clear at the outset that there are two alternative approaches to combining social sciences and quantum mechanics. One falls within the framework of physicalism, which postulates that the social sciences are derivable from first principles grounded in physics - an approach that is considered outdated and will not be discussed here - and the approach that Orrell[17]calls quantum-like, which instead only intends to employ the same mathematical concepts and tools derived from physics when these prove to be incidentally effective for this purpose, thanks to a favourable formal analogy, the ultimate reason for which is not the subject of further investigation.

The effectiveness of such mathematical tools is more due to the fact that the theory ultimately manipulates information since probabilities, information and observables are dealt with, rather than tangible physical entities.

This section outlines the salient concepts of quantum mechanics and illustrative cases in which these properties are recognisable in social systems and in particular in international relations, in order to introduce the habit of unconventional thinking through the recognition of patterns that can be traced back to the quantum scheme.

[17] Orrell, D. (2020). The value of value: A quantum approach to economics, security and international relations. *Security dialogue*, 51(5), 482-498.

2.1 QUANTUM THEORY, SCIENCE AND TECHNOLOGY

The first starting point for the creation of a common language substratum is to distinguish between three areas that are characterised and united by the adjective 'quantum', but at the same time are clearly distinct from each other. *Quantum theory* refers to that set of principles, such as the superposition principle, and mathematical rules - representing actual physical laws, such as the famous Schroedinger equation - that form the formal framework of quantum mechanics. Quantum theory is based on certain principles and certain quantitative inference rules that allow, above all, the calculation of probabilities of an entire spectrum of output possibilities, and the prediction of the future - in a probabilistic manner - from given initial conditions. Interestingly, there is actually not just one quantum theory, but many versions with increasing degrees of generality[18]. Generally, the one referred to is Bohr's original theory, which was later extended by Dirac and to versions with even more complex mathematical structures, so it is surprising how the simplest and thus also outdated version is still excellent for applications in quantum science and technology as discussed below.

The term *quantum sciences* refers to all those disciplines that extend their field of expertise through the inclusion of quantum methods or effects, ranging from quantum biology[19] to quantum decision theory[20], and on to quantum finance[21,22]. This is also where “quantum” diplomacy should be placed.

Quantum technologies ultimately consist of the engineering of processes based on nanometric objects that reveal quantum properties in order to isolate, transfer and process quantum information. Quantum technologies, which are discussed in Chapter II, have an application purpose. The aforementioned quantum information is based on the generalisation of bits (binary digits) of information into an equally coherent quantum version. The fundamental unit is the qubit (quantum bit), which is nothing more than a bit that - when queried - resists a random value that depends on the laws of quantum mechanics. Although this may be counterintuitive, this quantum type of bit enables theorists to invent a new type of algorithm and new types of protocols, which employ quantum properties including the “controlled randomness” of qubits.

---

[18] Prati, E. (2017). *Artificial Mind*. Ch. 3, EGEA Publisher.
[19] Lambert, N., Chen, Y. N., Cheng, Y. C., Li, C. M., Chen, G. Y., & Nori, F. (2013). Quantum biology. *Nature Physics*, *9*(1), 10-18.
[20] Yukalov, V. I. (2020). Evolutionary processes in quantum decision theory. *Entropy*, *22*(6), 681.
[21] Focardi, S., Fabozzi, F. J., & Mazza, D. (2020). Quantum Option Pricing and Quantum Finance. *The Journal of Derivatives*, *28*(1), 79-98.
[22] Agliardi, G., & Prati, E. (2022). Optimal tuning of quantum generative adversarial networks for multivariate distribution loading. *Quantum Reports*, *4*(1), 75-105.

2.2 THE APPLICATION OF THE CATEGORIES OF THOUGHT OF QUANTUM THEORY TO INTERNATIONAL RELATIONS

Without going into the mathematical and technical details of the theory of quantum mechanics, we can distinguish numerous examples of conceptual paradigm shifts from the previous history of science that this theory has brought about, and which have also been well formalised quantitatively. These conceptual categories include the discontinuous (being quantified), complementarity, indeterminacy, induction of system collapse following a measurement process, superposition, entanglement and non-commutation of operations. All these categories also find natural application in the study of international relations.

Contrary to Leibniz's belief that *natura non facit saltus*, at the nanoscale scale nature reveals instead that properties cannot be small at will but that quantities change in small jumps, called *quanta*, while intermediate values are forbidden. *Quantisation* disrupted the cornerstones of science a century ago and introduced conceptual novelties. In fact, quantisation is a concept familiar to everyone in everyday life, for example in the exchange of currency17. Everyone is accustomed to the idea that there is a minimum denomination in the currency, such as a cent of an euro or a US dollar, and that smaller quantities, such as half a cent, cannot be exchanged. In quantum theory, this also applies, for example, to the exchange of quantities of energy. This simple example shows how there is nothing revolutionary in itself in a concept that is the cornerstone of a scientific revolution, while the novelty lies in having imported science into a common concept in the economic and social spheres. In certain cases, as in this example, applying quantum mechanical concepts to the social sciences is actually a return to the origins where, however, the baggage with which one returns brings with it, analogous to the fictional Voyager 6 in the first Star Trek science fiction film, new knowledge derived from the confines of the universe - in this case borrowed from the dynamics of the very small.

Another of the concepts developed in the context of quantum mechanics is the principle of *complementarity*. It was enunciated by Niels Bohr in 1927 to address the paradox that light sometimes behaves as a wave and sometimes as a corpuscular particle. The solution was to recognise that both behaviours are possible but at the same time mutually exclusive, as they are also dependent on the nature of the type of measurement that is applied in the circumstance. To transfer this concept to the sphere of international relations, we can refer to the recent crisis in Ukraine. If Russia had proposed economic treaties to Ukraine, it would have been possible to know its willingness to cooperate economically, while military aggression on its soil made it possible to know its military response capability. Once the current military crisis scenario has been caused, it is no longer possible to know what the reaction from Ukraine would have been if Russia had proposed economic treaties. Knowing these two aspects is

alternative, since they cannot occur simultaneously and must therefore be considered complementary. Related to this is the concept of *indeterminacy*, which in quantum theory states that the accuracy with which a quantity is known affects (negatively) the possibility of knowing that of a complementary quantity, up to the limiting case in which perfect knowledge of one prevents knowledge of the complementary variable altogether.

A relevant category is that of system *collapse* when making a measurement, which is related to that of *quantum superposition*. In fact, quantum theory states that until the moment of measurement, the variable describing the system does not take on any precise value, whereas it is the act of measurement itself that makes the system collapse into one result rather than another. With collapse, only one of the potentials becomes an act. The peculiarity of quantum theory is that until the moment of measurement, assuming that the system can be in two alternative states, it maintains itself in a superposition of both these possible states. In this respect, quantum superposition, fundamental for example in quantum computing, is realised by particles possessing at the same time a mixture of different, even opposite properties: whereas a football can only rotate in one direction at a time, it is as if an atom could be made to rotate on itself in both directions at once. The two concepts of superposition and system collapse as a result of measurement are well suited to describe what happens when a population is asked to cast a vote, for example. Referring for example to the UK referendum in 2016, from the moment the referendum was allowed, the UK's future was twofold: remain in the EU or leave it. In quantum terms, the UK was in a potential overlap between the two future states, which did not become an act until the vote. The vote caused the collapse of the UK's potential future into one of only two possible states: the pro-Brexit front won and from that moment on the UK's future became decoupled from that with the EU.

To give another example of the effects of measurement as an act conditioning the system under observation, Wendt notes how language itself constitutes a measuring apparatus, which has an impact on what is observed. He states that,[23] "*in language what brings about a concept's collapse from potential meanings into an actual one is a speech act, which may be seen as a measurement that puts it into a context, with both other words and particular listeners"*. Measurement collapse begins with the decision to communicate one meaning rather than another, which in turn depends on the listener, whose understanding depends on how that language evokes his or her memory. Another example of the effect of measurement forcing the system to collapse into a precise value at the expense of all possible alternative potential values comes from finance. In that context, the price and volatility of a share can only be measured through the transactions themselves, which in turn alter these variables.

[23] Wendt, A. (2015). *Quantum mind and social science*. p. 217. Cambridge University Press.

Among the most peculiar phenomena of quantum mechanics is *entanglement*. It reveals how, at a fundamental level, particles that originally belonged to the same system remain connected to each other and the fate of one continues to cause consequences on the other as well. As an example of entanglement adapted to international relations, one can consider an ethnic group that is characterised by a sense of national identity but for historical reasons has been spread across the territory of different states, as was the case with the Kurds, who are spread across Turkey, Iran, Iraq and Syria, and in Europe. These communities, although spread over several states and also present in the West as a result of immigration, continue to remain connected and influence each other. When the persecution of the Kurds in Iraq took place in the 1980s, tens of thousands of Kurds sought refuge in Iran and Turkey, or Kurdish activists in the West joined the protests of Iranian Kurds over the death sentence and execution in 2020 of Kurdish activist Heidar Ghorbani in Iran.

Finally, let us consider the *non-commutativity* of transactions in international relations. We are generally accustomed to considering operations that switch: by changing the order in the sum of the shopping receipt, the result of the total does not change. In quantum theory, on the other hand, arrays of numbers are used instead of numbers, which is why the operations used do not commute. This seemingly unintuitive concept, however, can be represented very naturally in the field of international relations. If the operations being considered are, for example, an interference by a foreign power via a social medium targeting a party secretary, and the other is an election in the country, the fact that the order in which these take place is relevant is very obvious: a possible political scandal a few days before the vote has very different effects than if it is approached only after the election has already taken place. Consider, as examples, the role of chronological order in the scandal that hit Hillary Clinton in 2016 just days before the US elections, or the influence of Cambridge Analytics on the campaigns of several politicians, on Brexit in 2016 and on the elections in Mexico in 2018.

This set of conceptual categories, taken up by several authors in recent years, shows how quantum theory is general enough to be qualitatively adapted to international relations and offers perspectives and insights through their application.

3. CRITICAL CONSIDERATIONS ON THE QUALITATIVE QUANTIZATION OF INTERNATIONAL RELATIONS

Against the listed similarities and points of contact there are also cautious or sceptical considerations regarding the qualitative approach to international relations based on concepts derived from quantum mechanics. The first criticism is that this approach is based on affinities of a purely empirical nature. In other words, there is no intrinsic or theoretical reason why certain geopolitical phenomena or scenarios lend themselves to being described with concepts derived from the quantum realm. It is more of an observation of a state of affairs, which allows the characteristics of a scenario to be highlighted more consciously and described more accurately. This aspect highlights the circumscribed and non-universal nature of the qualitative application of quantum concepts to international relations. Consequently, there is no general theory whose specific cases decline on the scenarios of interest, but there are punctual correspondences between the space of possible scenarios that require description, and the categories revealed by quantum mechanics.

The second problem that emerges from this approach lies in the fact that the application of categories alone, without any quantification allowing for the formulation of models, exposes the impossibility of making predictions. It merely provides a change of perspective to bring out further understanding of a given scenario, making it possible, for example, to identify possible constraints, or relations between mutually exclusive quantities that are accessible from time to time. Forecasts, on the other hand, would possess a dual nature: informative in nature, since they allow something to be established about the future, and ontological in nature. In the latter respect, subsequent empirical investigation allows them to disprove or confirm the model. This is a fundamental aspect of modern science, i.e. checking both the verifiability and also the falsifiability of the model, as understood by Karl Popper[24].

As described in Chapter II, it is possible to further elaborate these qualitative models, and thus arrive at a quantitative phenomenology, based on mathematical laws describing social phenomena. For example, as in the case of quantum decision theory: it employs, instead of a probability logic based on Boolean elements linked by OR-type alternatives, an inclusive AND.

Awareness of this potential should not, however, lead one to overlook the importance that the application of the quantum categories mentioned above may reveal when it comes to assessing a geopolitical scenario. Firstly, when used as an analysis tool, they can support the understanding of a scenario or process,

[24] Karl R. Popper, *Science, Conjectures and Refutations*, in *Congetture e Confutazioni*, translated in Italian, Bologna, Il Mulino, pp. 68-69.

including the analysis of any limits to the knowledge that can be achieved. Consider in this respect the identification of complementary properties on a scenario, which implies that knowing one makes it impossible to know the other simultaneously. Secondly, they can offer insights towards the development of a corresponding quantitative theory, which will be based on a proper coding of the variables and processes involved. To conclude, quoting de Freitas and Sinclair[25]: "*We ask the reader to bear in mind, however, that any formal system will involve massive limitations and brutal simplifications. Probabilistic models of cognition are top-down – any such formalism will misrecognize much of the dynamic nature of events. But our aim is not to argue that quantum probability explains human judgment definitively – as that would go against the grain of the quantum – but rather to trouble reliance on classical probability and to invite speculation and experiment around different ways of reasoning with uncertainty*".

[25] de Freitas, E., & Sinclair, N. (2018). The quantum mind: Alternative ways of reasoning with uncertainty. *Canadian Journal of Science, Mathematics and Technology Education*, *18*(3), 271-283.

CONCLUSIONS:

- Quantum theory involves concepts, among which some of common use, that – thanks to their formal role in the theory – provide a ground for analysis in social sciences
- It is possible to use quantum theory to describe scenarios studied in international relations
- A model based on the qualitative use of quantum theory applied to a scenario of interest in international relations can facilitate the development of advanced quantitative models that calculate forecasts, such as quantum decision theory.

Chapter II

QUANTUM TECHNOLOGIES AND INTERNATIONAL RELATIONS

by Enrico Prati

SUMMARY: Quantum sensors (with metrology), quantum communications and quantum computers (with simulators), respectively generate, move under integrity and security constraints, and process by means of quantum algorithms, quantum data. United States first, and next other powers, have funded research towards such goals with several billions of dollars. USA are leaders in quantum computing while China – after taking advantage of western research at its beginning, in quantum communications. The expectations raised by such technologies should take into account the limitations that are currently addressed by ongoing research, but there are reasons not to fear a quantum winter after the current hype period. Quantum computers can be directly employed in the study of social sciences, and in particular in the framework of international relations. The case of the dynamics of terrorist networks in Middle East is reported as example.

1. TOWARDS INTEGRATIONS OF QUANTUM TECHNOLOGY

Over the last century, a series of technological revolutions have followed one another, so close together that there was no time for the development of one to finish when another was already in full acceleration. In addition to the impetus given by Maxwell's electromagnetism, these revolutions were underpinned by quantum mechanics and special relativity, which broadened the vision we had of the world, conditioned by everyday experience that was by its very nature incapable of perceiving phenomena in the very small and very fast, up to the speed of light. In that order, after radar and microwave technology, science produced nuclear technology, semiconductor and integrated circuit technology, lasers and telecommunications, which led to the development of computers and the Internet, hence the mature stage of artificial intelligence, all the way to nanotechnology. Quantum technologies are the product of these successive waves of innovation, thanks to nanometre-level control - practically at the atomic scale - of materials and fabrications, and advanced design tools with high computing power, which allow devices, sensors and processes to be simulated without the need to actually realise them until the design is addressed. Thanks to all these ingredients, and powerful theoretical models, it has been possible for some twenty years now to realistically think about manufacturing devices, instruments and apparatuses that fall under the name of quantum technologies. These are united by the fact that they encode or process information directly in objects that are generally nanostructured (with the exception of superconductors, which allow this even with a microstructure) to the extent that they *directly* exhibit the characterising properties of quantum mechanics listed in Chapter 1. Indeed, it is not enough for quantum mechanics to simply come into play, as has been the

case *indirectly* for the past seventy years for normal semiconductors. The qualifying and novel aspect consists in encoding or processing information by means of quantum states themselves, in a *direct* manner, whether based on single electrons, single atoms, by superconductors, or by quanta of light - called photons.

Quantum technologies fall into three large families: quantum communications, quantum computing (including simulations), and finally quantum sensors, into which metrology can also be subsumed.

Quantum sensors exploiting quantum mechanics can possess greater sensitivity, even below the noise threshold, and are also able to generate data already in quantum format, and thus compressed due to the quantum nature of the physical medium. There are gravity sensors and chemical molecule sensors, to give some examples, and there is talk of quantum radar and quantum imaging[26].

Quantum computers, starting with the foundational IBM and MIT conference "The Physics of Computation" held at MIT's Endicott House in Dedham, Massachusetts from 6 to 8 May 1981, were originally conceived to dissipate less energy by exploiting reversible processes, but with the discovery of quantum algorithms that disproportionately increased computing power - algorithms that can only be performed on quantum computers - interest has been turned almost exclusively to the resulting computing power[27]. There are problems that a quantum computer *can*[28] solve in minutes or hours that a normal computer would not solve in the lifetime of the universe (hence the term *quantum speed-up*).

Quantum communications, on the other hand, have both the purpose of ensuring the security and integrity of communication by employing the quantum states of light (quantum key distribution)[29], and also the purpose of transporting quantum data generated by sensors, or data leaving a quantum computer to reach a remote quantum computer (quantum Internet)[30].

Quantum technologies cover the entire data processing chain. Quantum sensors augment our senses and will contribute to a quantum Internet of things, quantum communications will be able to transport this data in an intact and secure manner, and quantum computers will process this data. The impetus for

---

[26] Lanzagorta, M. (2013, May). Amplification of radar and lidar signatures using quantum sensors. In *Active and Passive Signatures IV* (Vol. 8734, pp. 83-93). SPIE.

[27] Prati, E. (2017). *Artificial Mind*, Chapter 3. EGEA

[28] To date, a universal quantum computer of sufficient power to solve arbitrary problems with a systematic computational advantage is not yet available. This is why we are currently limited to convincing demonstrations of principle on a small scale.

[29] Cavaliere, F., Prati, E., Poti, L., Muhammad, I., & Catuogno, T. (2020). Secure quantum communication technologies and systems: From labs to markets. *Quantum Reports*, *2*(1), 80-106.

[30] Cuomo, D., Caleffi, M., Krsulich, K., Tramonto, F., Agliardi, G., Prati, E., & Cacciapuoti, A. S. (2021). Optimized compiler for Distributed Quantum Computing. *arXiv preprint arXiv:2112.14139*.

this wave of technology came with the two ARDA Roadmaps[31] on quantum computers and quantum communications in 2004. Today, many of the intuitions of twenty years ago have found their realization, realized. Nonetheless, we must be cautious when assessing their impact and robustness. Each sensor, device, circuit and apparatus are taken to its technical limit and operates under extreme conditions in terms of sensitivity and exposure to interference. Since quantum states are difficult to isolate and are very fragile, and therefore deteriorate quickly according to the process of decoherence[32,] the engineering problems are much more relevant than the conceptual ones. Therefore, it is premature to expect commercial success to date, but at the same time all governments have launched investment plans to support the technology transfer of these technologies to create value in the industrial fabric, either by incorporating them into mainstream technologies or by creating radically innovative systems. Evolved venture capital funds also contribute to these, enabling “deep tech” start-ups to develop products based on quantum technologies, exposing themselves to great risk against potential great rewards.

[31] Bennett, C. H., & Brassard, G. (2004). A Quantum Information Science and Technology Roadmap. *Part*, *2*, 12. ARDA - LA-UR-04-4085

[32] Porotti, R., Tamascelli, D., Restelli, M., & Prati, E. (2019). Coherent transport of quantum states by deep reinforcement learning. Nature *Communications Physics*, *2*(1), 1-9.

2. GEOPOLITICS AND QUANTUM TECHNOLOGIES

The country that pioneered and invested the most in quantum technologies is the United States. On the strength of a forward-looking and meritocratic public funding system for research, it started scouting for these technologies and defined a technology roadmap in 2004, later updated in 2007, which stimulated research mainly in public centres.

Subsequently, the funding turned to business development and over a period of about ten years resulted in several industrial projects for the development of a quantum computer by public companies, such as Intel, IBM, Google and Microsoft. The order of magnitude of the funding provided by these giants of the North American information technology industry is between USD 200 million and USD 500 million for each of them, each with many dozens - even hundreds - of people involved in the engineering of all levels, from the hardware layer up to the software applications. Canada has also provided significant funding and led to the creation of the quantum computer DWave, which is based on an alternative architecture called quantum annealing that was first proposed by two Italian scientists - Bruno Apolloni and Diego De Falco in 1988. Recently, at the opening of the annual High Performance Computing and Quantum Computing Workshop organised by the author in collaboration with CINECA[33], the two scientists recalled how at the time of their pioneering studies the number of experts in the world dealing with this idea numbered just a few, whereas today, quantum computing includes NASDAQ-listed companies such as the US-based IonQ and is taught in Master's degree courses.

In response, China has also promoted the development of quantum computing projects (Baidu and Alibaba) based on the strengthening of public research initiatives, such as that of the Chinese Academy of Science (CAS). Chinese quantum computing scientists are a textbook example of China's multi-decade scale planning and adequately resourced talent return programme. China has brought back PhD students and post-docs visiting the US temporarily recruited from quantum computing research centres and provided them with university professorships and substantial funds to purchase state-of-the-art equipment. The same talent return operation also earned China the lead in quantum communications. After completing his doctorate at the University of Vienna with Prof. Zeilinger, universally recognised as one of the founders of quantum information, Jian-Wei Pan moved to the University of Heidelberg in Germany where he received financial support from the European Union, including a

[33] D. Ottaviani, R. Mengoni and E. Prati, IV High Performance Computing and Quantum Computing Workshop, 15-16 December 2021, online

Starting Investigator Research Grant from the European Research Council from 2008 to 2013 amounting to almost 1.5 Meur[34]. He was then recruited by the University of Science and Technology of China (USTC), where in 2016 he led the project to launch the first Micius satellite of the Quantum Experiments at Space Scale, which in 2017 demonstrated earth-space coupling via quantum communication[35] on a scale of 1200 km. On the strength of these successes, China also entrusted him with the construction of a photon-based quantum computer called Zuchongzhi 2.1, which was claimed to be a million times faster than Google's Sycamore chip. Jian-Wei Pan continues to lead a group at the University of Heidelberg made up of 80 per cent Chinese personnel, but the earth-space communication network that has achieved supremacy is built in China, which is funding quantum technology research with $15 billion. Anhui Province alone has launched a USD 1.6 billion Quantum Science Industry Development Fund.

There are a large number of excellent Italian scientists in quantum technologies abroad, both in the academic and private sector, but the rate of return is negligible. In a logic of national interest, to make use of the skills of Italians in the world, it would be necessary to set up university professorships of indefinite duration, clothed with multi-year funding and linked to commensurate dedicated spaces on the public side, and to strengthen funding measures for new start-ups on the private side. In this sense, scientific poles such as the Human Technopole in Milan and the new High Performance Computing centre being set up in Bologna possess the characteristics that would be needed to support such initiatives.

Different countries have approached the opportunities offered by quantum technologies with different attitudes. A measure of the results of these policies can be read through the number of patents filed in the various fields. A distinct growth trend emerges in US patents in the area of quantum computers, while the trend in quantum communications in China is equally distinct (Figure 1). The European Union has approved a EUR 1 billion flagship programme over 10 years among 25 countries of the Union, thus averaging about EUR 4 million per country per year, which cannot replace the effects of national programmes such as those put in place by some countries. To give an example, the Waterloo region in Canada alone has put in place USD 568 million for the Institute for Quantum Computing, USD 205 million in Quantum Valley Investments, and USD 591 million in the Perimeter Institute, for a total of more than USD 1.3 billion. The results have been the development of several quantum computing companies such as DWave and Xanadu for hardware, or Zapata and 1Qbit for software, or

[34] https://cordis.europa.eu/project/id/202499/it

[35] Yin, J., Cao, Y., Li, Y. H., Liao, S. K., Zhang, L., Ren, J. G., ... & Pan, J. W. (2017). Satellite-based entanglement distribution over 1200 kilometers. *Science*, *356*(6343), 1140-1144.

even quantum communication companies such as EvolutionQ, which are already worth more than the funding provided. The investment therefore generated value and contributed to geopolitical supremacy as quantum-haves over quantum-haves-not.

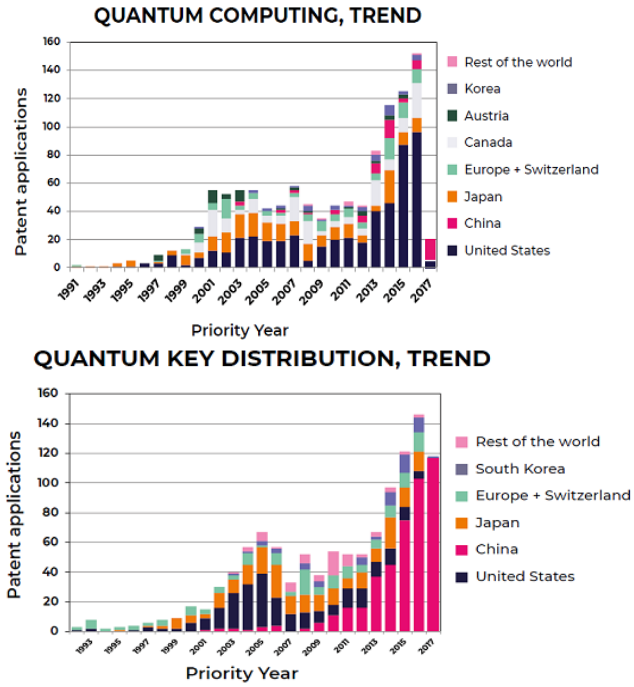


Figure 1: Number of patents repeatedly in quantum computing and quantum key distribution, by country up to 2017. There has been a slow move away from quantum key distribution (QKD) by the US, which has focused on quantum computers, while China has invested mainly in such quantum communications. Recently, China has also reused QKD technologies for quantum computers. Japan has produced more patents than the whole of Europe. Source: European Commission Roadmap on Quantum Technologies (2018).

Russia too, with its USD 790 million funding to the Russian Quantum Centre, has decided to make public its commitment to quantum competition[36]. Achieving independence in terms of quantum technologies once they have reached maturity will primarily lead to an advantage over other countries in terms of utilisation, but also to a quantum divide, between those countries that are part of the circle of those with superior computational power, greater awareness and surveillance, and more secure communications and those that are dependent on them or are excluded from them altogether. Possessing such technologies also allows in-depth knowledge in terms of its vulnerabilities. Indeed, it is necessary to be

[36] https://www.nature.com/articles/d41586-019-03855-z

aware that while quantum communications promise an inviolable and guaranteed method as a system operating *under idealised conditions*, it is equally true that practical realisations have to reckon with the whole component of signal management electronics and photonics, which has enabled at least 30 different known attacks on both transmitters and receivers over the last twenty years. In other words, the scientific community has made public a new type of QKD attack every eight months for two decades, not counting unpublished ones. Several observers recommend caution and there are agencies that deprecate - at present - the use of QKD methods for encoding vital messages.

In conclusion, it is to be expected that quantum technologies will play a role similar to that of artificial intelligence, telecommunications and semiconductors in terms of geopolitical leverage, both in terms of the economic spin-offs it confers and the technological supremacy it will bring in terms of information control and processing. There are, however, still important steps to be taken in each of them, including the technological ones to arrive at robust products capable of facing the market, and also the certification one to make these prototypes also marketable products. In this respect, there are standardisation initiatives such as the one promoted by NIST and IEEE in the USA, and certification initiatives such as the one promoted by ETSI, of high potential impact, which can also decree competitive advantages and disadvantages in the development of future quantum technologies. For instance, there are currently two initiatives to standardise the security certification of ISO/EN 15408 “Common Criteria” QKD systems such as the ETSI ISG-QKD and the ISO SC27 WG3, which is in the process of publishing these standards. Standards development and certification communities are open and, as such, penetrable, thus being susceptible to any interests brought by the contributing actors.

3. QUANTUM COMPUTERS: HYPE OR TECHNOLOGY PLATFORM FOR SOCIAL SCIENCES?

Quantum computers represent a potential breakthrough in the field of computation since, thanks to the use of quantum bits (*the qubits*), it is possible to devise and apply quantum algorithms that are able to reduce even exponentially the number of steps to be taken to arrive at the solution of certain problems[37]. The applications explored range from finance, renewable energy[38], logistics, the design of new materials for avionics, the chemistry of more efficient batteries, distribution in networks whether in the energy, road traffic or telecommunications sectors, and last but not least, artificial intelligence in its quantum version[39].

However, while there is no doubt that realising quantum computing is a reality - think of the quantum computers in the cloud made available by IBM, DWave, Rigetti and IonQ to name but a few - there is debate as to the time perspective within which quantum computers will be able to accommodate problems of sufficient size to give a real advantage over traditional computers.

An important aspect of evaluation is that there are many different types of quantum computers. Quantum computers follow four possible alternatives of computational architecture (circuital, adiabatic, one-way and topological) that differ in how data are encoded and then processed. At the same time, they can be realised in practice with different technologies, such as superconductors, atoms in vacuum, semiconductors and photons, each of which may or may not be suitable for the various architectures mentioned above, so that only certain combinations of architecture and technology are actually possible. Since 2017, the number of qubits handled in a quantum chip has been growing steadily year on year in all the technologies mentioned up to the current hundred (thousands in the case of adiabatic architecture). Although a comparison between technologies is not easy, every year there can be a flip-flop in estimating which hardware technologies have the most computational resources, marking a clear trend that resembles the trend of miniaturisation of transistors (Moore's law) in the early 1960s. Despite the fact that this quantum hardware is in fact commercialised prototypes, industries have started experiments on use cases in order not to be unprepared should quantum computers mature in their development and offer greater computing power than is already the case with conventional computing

[37] Shor, P. W. (1999). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, *41*(2), 303-332.

[38] Giani, A., & Eldredge, Z. (2021). Quantum computing opportunities in renewable energy. *SN Computer Science*, *2*(5), 1-15.

[39] Rocutto, L., Destri, C., & Prati, E. (2021). Quantum semantic learning by reverse annealing of an adiabatic quantum computer. *Advanced Quantum Technologies*, *4*(2), 2000133.

methods. A large company not taking an interest in this innovation could result in its exclusion from the market in the not-too-distant future, similar to what happened after the arrival of Netflix to the well-known film rental chain Blockbuster, which used to preside over all the cities in the West and disappeared into thin air. Since, however, the number of resources in terms of qubits in a quantum processor is of the order of 10-150 qubits, one can currently only test small-scale, principle demonstrations on very small use cases. One thinks, for example, that for certain algorithms, millions of qubits may be needed to ensure that they can process the calculation, and at the same time carry out the necessary quantum error correction. In fact, it should not be forgotten that the consumption of qubits is exacerbated by the fact that in order to protect the information of a qubit from environmental disturbances and keep it alive for the duration of any calculation, an average of 100 to 1000 additional qubits are needed.

It is therefore legitimate to expect that, despite major funding at global level and the involvement of important early users (such as Airbus, JP Morgan, General Electrics, Volkwagen, and DHL, to name but a few, or Banca Intesa and ENI to remain in Italy), the excessive protraction of a situation of developing commercial prototypes without yet arriving at a clear advantage could lead industry to become disillusioned with the feasibility of the universal quantum computer, resulting in a quantum winter.

However, even taking into account the principle of caution that one must maintain when dealing with emerging technologies, there is an argument not to prefigure for the current scenario a future similar to the so-called AI winter, which preceded the current era of full maturity of artificial intelligence by two decades. Quantum algorithm pioneer Umesh Vazirani, for example, is convinced that unlike artificial intelligence, which suffered from conceptual bottlenecks in the 1970s that were not yet understood and have only recently been overcome, there is no such bottleneck with quantum computers. This is, for example, because two different conditions apply. The first is that there are many competing technologies based on completely different technologies that back each other up. The second is that there are no conceptual problems in the theory, which is fully understood and can only increase in terms of further improvements and contributions, but rather engineering issues that have been addressed for a relatively short time and with still small and growing communities.

That said, it is useful to examine in greater detail whether it is possible, among other applications, to apply the power of quantum computers to the social sciences and thus to international relations as a use case of choice.

The assumption is firstly that there are quantitative models, in which the quantum categories applied to describe a scenario of interest in international relations are employed by means of mathematical equations derived from quantum theory.

The plausible scenario is to regard quantum computers as a promising exploratory tool to be used alongside classical computers, while waiting for them to reach full maturity and be able to contain quantum problems large enough not to be solved any other way or so quickly.

Methods suitable for quantum computers to perform community detection are generally based on the already studied methods including traditional, dynamic, local and overlapping detection methods. For example, studies have been conducted on quantum social networks (QSNs), which have proven to be more efficient than tools based on classical networks on specific problems[40]. With the quantum many-body theory, mathematical models based on the complexity theory have been developed and validated through the simulation of social networks, performing social network analysis (SNA), which leads to insight into the dynamics of the social network under investigation[41]. Again, studies of the evolution of the *entropy* of a social network propose to study quantum consensus systems to establish a relationship between the quantum component versus the classical component describing the network, even to the point of considering quantum gossiping.[42] Akbar and Saritha collected and described some 20 methods of social network analysis employing algorithms for quantum computers[43].

There is also a hybrid area between artificial intelligence and quantum computing, which consists of quantum artificial intelligence and its quantum machine learning branch. Since artificial intelligence is able to reconstruct from incomplete data, identify patterns, predict time series, similarly quantum machine learning algorithms[39,44,45] can in principle also be applied to studies concerning international relations where conventional machine learning could be applied.

In the following section, we will examine a specific example case of how to use a quantum computer for the analysis of a geopolitical scenario.

---

[40] Cabello, A., Danielsen, L. E., López-Tarrida, A. J., & Portillo, J. R. (2012). Quantum social networks. *Journal of Physics A: Mathematical and Theoretical*, *45*(28), 285101.

[41] Bisconti, C., Corallo, A., De Maggio, M., Grippa, F., & Totaro, S. (2010). Quantum modeling of social dynamics. *International Journal of Knowledge Society Research (IJKSR)*, *1*(1), 1-11.

[42] Fu, F., Christakis, N. A., & Fowler, J. H. (2017). Dueling biological and social contagions. *Scientific reports*, *7*(1), 1-9.

[43] Akbar, S., & Saritha, S. K. (2020). Towards quantum computing based community detection. *Computer Science Review*, *38*, 100313.

[44] Lazzarin, M., Galli, D. E., & Prati, E. (2022). Multi-class quantum classifiers with tensor network circuits for quantum phase recognition. *Physics Letters A*, 128056.

[45] Agliardi, G., & Prati, E. (2022). Optimal tuning of quantum generative adversarial networks for multivariate distribution loading. *Quantum Reports*, *4*(1), 75-105.

4. A CASE STUDY: THE EVOLUTION OF GLOBAL TERRORIST NETWORKS

In the year 2017, at the Los Alamos laboratories in the USA, the three researchers Ambrosiano, Roberts and Sims published the technical report LA-UR-17-23946 commissioned by the Department of Energy, in which they investigated for the first time the application of a quantum computer to a social science study based on a model describing an equilibrium of geopolitical interest[46]. This study employs the 2000-qubit Dwave quantum computer and actually implements a problem that goes by the technical name of bipartition of a graph. The model underlying the use of the quantum computer is based on the following assumptions: let us assume that we can describe a social network by means of connections between the elements of that network (called the *nodes*) that correspond to friendship and enmity relations (positive and negative connections, respectively). The fundamental question of the structural balance of this network is whether it is balanced or not. Such a social network is balanced if the network of friendship or enmity connections allows it to be bipartite, i.e. divided between only 2 factions, within which only friendship relations exist, and between which only hostile relations exist. The less the connections respect this simple rule, the more unbalanced this network will be, making the two factions more ambiguous. Sociology suggests that such unbalanced networks are unstable and associated with more violence[47].

There is an analytical method for assessing the degree of balance in social networks, which is done by distributing the various subjects, each of which is associated with a node in the network, among the factions in such a way that within each faction there is a minimum, possibly zero, number of hostile relationships. This problem, from a mathematical and computational point of view, is of the highest degree of complexity (called NP-hard). This is where the possibility of employing a quantum computer comes in. In fact, this type of search - i.e. assigning groups to factions while introducing as few nodes as possible linked by a hostile relationship - is mathematically equivalent to searching for the minimum energy of the quantum circuit describing the set of friendship and hostility relationships. This type of search for the minimum energy of a quantum circuit is effectively performed by an adiabatic quantum computer, such as the D-Wave 2X based on superconducting qubits.

---

[46] Ambrosiano, J. J., Roberts, R. M., & Sims, B. H. (2017). *Using the D-Wave 2X quantum computer to explore the formation of global terrorist networks*. Technical report LA-UR-17-23946, Los Alamos National Laboratory.

[47] Nakamura K., Tita G., Krackhardt D., "Violence in the 'Balance': A Structural Analysis of How Rivals, Allies, and Third-Parties Shape Inter-Gang Violence", Heinz College Research, ResearchShowcase@CMU, http://repository.cmu.edu/cgi/viewcontent.cgi?article=1411&context=heinzworks, (2011).

The problem submitted to the quantum computer was therefore to assess the deviation from the perfect equilibrium of two scenarios characterised by the strong presence of militant organisations in two theatres in particular, Iraq and Syria. Specifically, the data collected by Stanford's Mapping Militants Project in the years from 2000 to 2016 were examined for both scenarios, yielding networks of the order of twenty nodes, equal to the number of militant organisations that are in the order of 20-30, and associating with each pair the information as to whether they were friendly or hostile.

Compared to the results observed in this study, the degree of balancing began to decline in the Syrian theatre as the Islamic State entered a context already populated by other groups. It could also be observed quantitatively that while the social network grew over a given period, the imbalance associated on average with each individual node did not vary significantly, a fact that can be explained by adaptive behaviour of the various factions.

It should be pointed out that for this problem, on this scale of actors involved, there is no advantage over the answer obtained by ordinary computational means to the same question and should therefore be considered as a demonstration of principle. However, since this is a computationally hard problem, the calculation time will increase exponentially as the number of nodes (of factions) increases, whereas for the quantum computer, a single evaluation will still suffice, resulting in an extremely advantageous calculation time. At the time of this study, essentially all 1150 qubits of the DWave quantum computer available at the time were used. Five years on, the number of qubits has more than quadrupled and the number of connections of each qubit with the others has doubled, showing how rapidly the field is evolving.

CONCLUSIONS:

- Quantum technologies include sensors, communications and computation
- Quantum technologies differ in their degree of maturity and are all under development, including commercialised prototypes
- In particular, quantum computers will be a lever in the balance of power between states
- The risk of a technology hype for quantum computers is mitigated by the variety of the possible technologies to build it, and the absence of conceptual bottlenecks
- As a general-purpose technology, quantum computers can be applied to international relations, given a quantitative model based on quantum theory
- The quantum computer DWave has been used to study the formation of terrorist networks in Syria and Iraq.

Chapter III

THE USE OF ARTIFICIAL INTELLIGENCE IN WEAPONS SYSTEMS: INTERNATIONAL LEGAL FRAMEWORK

by Andrea Gilli and Lucrezia Scaglioli

SUMMARY: The technological acceleration of the past few decades in processors, big data and machine learning has led to the revolutionary progress of artificial intelligence. These developments have renewed the attention and increased the pessimism regarding the potential risks coming from the use of this technology, especially in the military domain. For this reason, scholars and experts started requiring new and increased regulation, control and even the banning of this technology. However, we consider it is worth investigating whether these widespread fears are legitimate and whether artificial intelligence can endanger more risks than opportunities. In this chapter, we will attempt to answer this question, starting from the origins and developments of artificial intelligence, the scope and implications of this new technology applied to the military domain and autonomous weapon systems. Then, on the basis of this analysis, we evaluate the criticism, by experts and scholars, of the risks associated with its use. Finally, we look at the ethical and legal framework which defines the international rules for verification and use.

1. ARTIFICIAL INTELLIGENCE: DEVELOPMENT AND SCEPTICISM

The developments and acceleration of artificial intelligence since the 2000s have been accompanied by an equal wave of criticism and pessimism regarding its use. In particular, major concerns relate to the use of AI applied to the military domain. With the advent of new autonomous weapon systems, some experts and scholars, in fact, argue that we are facing a new military revolution capable of changing the technology, redistributing military power, causing new, more frequent and more lethal conflicts, and bringing about upheavals in the world balance. Before such predictions can be verified, however, it is necessary to understand the origins, development methods, functions and future growth prospects of artificial intelligence. In fact, the widespread pessimism towards AI is based on three generic assertions: its continuous and increasing progress, the predominance of the commercial sector and thus its increased diffusion, and finally, the great pervasiveness of AI. By analysing the developments, the technical-technological articulation and the possible military uses of AI, it will be possible to understand the limits of both technology and of this scepticism, the focus of which is solely on consequences. Part of the criticism levelled at AI, finally, concerns the field of ethics and international law. The last part of this chapter will therefore analyse the ethical principles that future AI developments will have to comply with, in addition to the legal framework already in place.

2. TECHNOLOGICAL REVOLUTION

In this section, we will deal with the technological development that has led to the immense acceleration of artificial intelligence, giving a definition of what AI is, what functions it can have, and finally, what kind of technological structure it needs to continue its expansion. We will then describe the two types of approaches to AI, from Good Old Fashion AI (GOFAI) to Deep Learning (DL), the shift of which is due to the exponential progress in the field of processors, big data and machine learning.

Developments and continuous advances in technology have led to three different industrial revolutions in human history, social and political changes.[48] From the increase in computation, and consequently in precision levels, to the advent and development of new electronic and computer technologies, from artificial intelligence (AI) to machine learning (ML) - the computational key to artificial intelligence - and big data (BD), we are facing what some, like Karl Schwab, call the Fourth Industrial Revolution or others, the Second Machine Age.[49] This era is characterised by an exponential increase in the development and diffusion of technology, more pervasive and faster than previous waves of evolution, leading to major transformations, touching every aspect of our lives, with socio-economic and political-international implications.[50] Paradoxically, the pandemic itself has led to a further acceleration towards this transformation and digitisation of our lives, forcing remote work and the use of digital technologies. In fact, computational power has increased exponentially since 1965 due to the development of processors, algorithms and data.

Artificial intelligence, which lies at the heart of this revolutionary new wave of technology, is defined by some as general-purpose technology (GPT), but more generally it can be interpreted as a type of technology aimed at simulating the intelligence of human beings and whose impact, it is already possible to predict, will occur in multiple fields, from the economic to the military. Artificial intelligence can be applied to different domains for different functions: it can be used to direct physical objects, such as robots, without human control, or to process and

---

[48] Headrick D. R. (2010), *Power over people: technology, environments and Western imperialism, 1400 to the Present*, Princeton University Press, Princeton, NJ; Onorato M., Scheve K. and Stasavage D. (2014), "Technology and the era of mass army", *The Journal of Economic History*, Vol. 74, No. 2, pp. 49-81; Boulanin V. and Verbruggen M. (2017), *Mapping the development of autonomy in weapon systems*, SIPRI, Stockholm.

[49] Schwab K. (2016), *The fourth industrial revolution*, Crown Business, New York; Brynjolfsson E. and McAfee A. (2014), *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies,* New York, NY: W. W. Norton & Company.

[50] Horowitz M. C. (2018), *Artificial Intelligence, international competition and the balance of power*, Texas National Security Review, Vol. 1, No. 3, pp. 37-57.

interpret information, or finally, through the overlapping of several specific functions, it can be used for new forms of command and control (C2).[51] There is no common agreement among scholars as to what the predetermined scope is or should be. However, it is generally agreed that the impact, precisely because of its pervasive nature, will be such that it will bring change, transformation and renewed competition to the international political scene, often compared to the effects of electricity or the steam engine before that. As with processors, there are two types of artificial intelligence. One is defined as general, i.e. capable of performing multiple functions in parallel, representing the type of AI that in a future scenario would be able to replace human beings. The other model, on the other hand, has a more limited and specialised approach to precise fields of action, thus defined as narrow.[52]

There are two types of meta-approaches to AI: top-down and bottom-up. The former, also referred to as Good Old Fashion AI (GOFAI), is based on a deductive approach, whereby all information must be coded and entered *ex ante*. This was the predominant approach until the 2010s and precisely because of its structure, which implies a theoretical codification of every possible scenario, its clear limitations and shortcomings are evident. With the exponential acceleration of semiconductors, chips, processors and algorithms, machine learning techniques have undergone new developments. To quantify these transformations, one only has to think of the increase in data production, from five exabytes in 2003 to 59 zettabytes - or 59 trillion gigabytes - in 2020, and the rapid decrease in the cost of 3D Lidar*(Light detection and ranging*) sensors, from USD 30,000 in 2009 to USD 80 in 2019.[53] Calculation sequences that would have taken 89 years in 1982 are now solved in seconds.[54] This improvement was also made possible by the increase and specialisation of processors. There are two types of microprocessors: specialised microprocessors, i.e. used for specific functions and sequences, and general-purpose microprocessors that are capable, on the other hand, of being used for multiple, parallel applications.[55] Specialised processors have started to be demanded and developed in increasing numbers, especially since 2010, given their fundamental role in sequential operations in the operation of machine learning algorithms. Developments in algorithms, in turn, are the consequence of new ML techniques

[51] Horowitz M. C. (2018), *Artificial Intelligence, international competition and the balance of power*, Texas National Security Review, Vol. 1, No. 3, pp. 37-57.

[52] See Russel S. and Norvig P. (2010), *Artificial intelligence: a modern approach*, Upper Saddle River, NJ, Prentice Hall, 3rd edition.

[53] Lee K. F. (2018), *AI Superpowers: China, Silicon Valley and the new world order*, Boston, MA, Houghton Mifflin.

[54] Brynjolfsson E. and McAfee A. (2014), *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies,* New York, NY: W. W. Norton & Company.

[55] Henessey J. L., Patterson D. A. (2019), *Computer architecture: a quantitative approach*, Morgan Kaufmann, Sixth Edition, Cambridge, MA.

and the expansion in the use and functions of software, including its application to artificial intelligence systems.[56] This progress was also made possible by the increase in the production and availability of data. With the digitisation of different types of information and the spread of portable devices such as phones, laptops, tablets, etc., the amount of data available has become exorbitant.[57] These three enhancements in microprocessors, ML techniques and available data have led to renewed attention, new investments in the field of artificial intelligence and the move from GOFAI to the inductive, or bottom-up, approach based on deep learning, i.e. letting the AI learn and improve thanks to the patterns, trends and prediction capabilities derived from the huge amounts of input data and its interactions with the world.[58] It is a concept that, as such, has existed since 1965, introduced at the conference at Dartmouth College in Hanover, New Hampshire, but which has experienced an acceleration in its development and use particularly since 2010 onwards based on the exploitation of deep learning, for the reasons stated above.[59]

3. MILITARY IMPLICATIONS

Having mentioned the origins and developments of AI, it is appropriate to turn to its implementation in the military field. The consequences of these transformations for the military, defence and security domain are still at an early stage of development and the implications of their use are not entirely clear. Certainly, however, these new technologies pose new opportunities and advantages, as well as questions, challenges, risks and concerns.[60] The defence and security sector have, for several years now, begun the process of integrating AI into its force structure. However, the debate often focuses solely on the consequences, raising concerns and criticism. In fact, there is talk of an arms race, of the risks due to the increased spread and pervasiveness of autonomous weapon systems, and even of the upsetting of the balance of power. In this

[56] Allen G. (2020), *Understanding AI technology*, US Department of Defence, Joint Artificial Intelligence Center, Washington, DC.
[57] Manyika J. Et al. (2011), *Big data: the next frontier for innovation, competition and productivity*, McKinsey Global Institute, New York.
[58] Gilli A. and Gilli M. (2021), 'Technology and new global conflicts', *Aspenia*, Vol. 94, pp. 117-127.
[59] Haenlein H. and Kaplan A. (2019), "A brief history of artificial intelligence: on the past, present and future of artificial intelligence", *California Management Review*, Vol.61, No.4.
[60] Horowitz M. C. (2018), *Artificial Intelligence, international competition and the balance of power*, Texas National Security Review, Vol. 1, No. 3, pp. 37-57.

section we will touch on these points and analyse whether or not these fears and concerns are justified.[61]

The expression used by journalists, politicians and researchers of an “artificial intelligence arms race”, apart from increasing a debate of strategic competition, does not coincide with factual reality. Major powers are indeed competing for research and development of military and commercial technologies related to artificial intelligence, but this does not fit the traditional definition of an “arms race”.[62] Although the phenomenon does not exactly fit the definition of an arms race, the acceleration in developments and potential uses of artificial intelligence in the military field are being emphasised by political leaders, CEOs and academics as a true military revolution. Examples are the Chinese government's 2017 goal of achieving world hegemony in the field of artificial intelligence, the introduction of military strategies for AI by major European powers and NATO, the Pentagon's Project Maven programme, or Putin's statements, for whom controlling artificial intelligence entails world domination.[63]

Artificial intelligence, we recall, aims to mimic human behaviour and reasoning through a chain of information from perception to cognition and finally to action. This functioning means that autonomous AI systems determine their actions on the basis of reasoning and probabilistic calculations determined by sensor inputs, which must perceive the surrounding world and deconstruct it, but still without logical connections.[64] What for a computer requires enormous amounts of data to process and some sort of training is done in seconds by the human brain, demonstrating the paradox of this so-called fourth industrial revolution. The more articulated, developed and complicated the technologies are, the more there is, and will be, a need for humans to control, direct and interpret them.[65]

The military implications of AI are manifold and have various advantages for the defence and security domain. Firstly, AI enables greater data mining, collection, transmission and analysis, thanks to radar and sensor improvements, enhancing intelligence, surveillance and reconnaissance actions. Secondly, it makes it possible to optimise combinatorial problems and thus improve logistics with the use of unmanned autonomous vehicles by land, air or sea. Thirdly, it increases precision in targeting the enemy through precision-guided weapons, the implication of which can be seen as a reduction in collateral damage to civilians. It also speeds up war time, consequently increasing predictive analysis

---

[61] Frey B. C. (2019), *The technology trap: capital, labor, and power in the age of automation*, Princeton, Princeton University Press.
[62] Scharre P. (2021), *Debunking the AI arms race theory*, Texas National Security Review, Vol. 4, Issue 3.
[63] Gilli A. and Gilli M. (2021), 'Technology and new global conflicts', *Aspenia*, Vol. 94, pp. 117-127.
[64] Cummings M. L. (2017), *Artificial Intelligence and the Future Warfare*, Chatham House.
[65] Ibid.; Gilli A. (2020), *“NATO-Mation”: Strategies for leading in the age of Artificial Intelligence,* NDC Research Paper N. 15, Rome, NATO Defence College.

and decision-making actions, which AI can perform faster than a human brain.[66] There are further aspects to be taken into account:

- The use of AI in the military sphere further reduces the relationship between labour and capital, a trend that began with the secular industrial revolution.
- Whereas in the first industrial revolution machines replaced the energy produced by the muscles of humans, animals or nature, in this industrial revolution machines seek to replace the cognitive capacities of individuals. At this stage, however, algorithms and robotics are still mainly used for 4D (Dangerous, Dull, Dirty, Dumb) missions and the limits of deep learning cannot yet allow for uses that fully replace beings.
- Exploiting the AI allows for superior speed and accuracy thanks to algorithms that speed up attack times and discriminate enemy targets more accurately. Reaching speeds beyond human capabilities.
- However, the use of AI in the military field also introduces a command-and-control problem, related to Human Machine Interaction (HMI) or Human Machine Teaming (HMT). AI generates a question of trust and reliability between the (military) commander and the robots or automated systems. There is also the question of how to train and instruct humans to work better with these systems, such as designing interfaces, understanding human psychology in interacting with autonomous and automated systems, or reviewing training.
- Finally, the industrial base must be considered. If in the past the industrial base was essential to win wars, in the post-industrial era, characterised by software and big data, it will be necessary to adapt the industrial base to this era to ensure future security.

Despite the incredible developments in the field of autonomous systems, whether by air, land or sea, the transition to actual implementation in military operations is still far off, partly as a consequence of the necessary organisational and structural adaptations, the enormous costs, partly due to the priority given to the development of traditional vehicles and weapon systems, and finally, according to some, because at this stage the driving sector in the development and use of UVs, such as drones or driverless cars, is the commercial and private sector.[67]

The developments and integration of autonomous weapon systems has raised widespread debate and growing concerns about the beginning of the robotic age and the diminishing human control in the military. The most common concerns of scholars and experts range from the risk of increased instability and conflict to the disruption of the international order and balance. Criticism generally rests on

[66] Ibid.
[67] Cummings M. L. (2017), *Artificial Intelligence and the Future Warfare*, Chatham House.

three main assumptions: a greater spread of autonomous systems in the years to come, a wider pervasiveness of these systems, and a continuous advancement of AI. In this section, moving from the criticism, we analyse the risks that AI may or may not entail.

The first concern about AI developments, driven mainly by the commercial sector, is its more rapid and pervasive deployment for two reasons: the lower unit price of production and the easier dissemination of these by private individuals incentivised by profit and economies of scale. Greater propagation would, on the one hand, limit the strategic military advantages and, on the other hand, increase and expand the pervasiveness of these technologies. The risk interlinked with this possibility would mean that more actors could have easy access to both commercial AI systems and lethal weapon systems, leading to increased conflict and international political instability.[68] this type of criticism, however, is based on the assumption that AI is a cheaper and easier technology to implement, replicate and disseminate than traditional weapon systems. In fact, while it must be acknowledged that commercial technology is relatively cheaper, it must be specified that once moved to the military sphere, that technology needs increasingly specific and costly requirements for fewer units, which prevents the exploitation of economies of scale.[69]

A second type of criticism concerns the pervasiveness and acceleration in the use of AI. This increased spread could bring about changes to the nature of warfare: making it potentially faster, more unstable and more lethal. This is another reason why lethal automated weapon systems, or LAWs, create apprehension and aversion among both the public and the military itself.[70] Easier to produce, easier to imitate and cheaper, according to some scholars, drones, for instance, could change the dynamics of international politics by redistributing military power, resulting in increased instability and the frequency of new, faster and more deadly conflicts. Indeed, the rise in the use of robots incorporating artificial intelligence, or autonomous weapons - Unmanned Autonomous Vehicles (UAVs) - while being hailed as the first stage towards a new technological-robotic era, also increases the concern and debate of the last 15 years about the possibility of banning what are referred to as “killer robots” that would lead to crises, violence and human rights violations, for instance due to a malfunctioning or inaccuracy of the algorithm.[71] However, even this criticism can be seen as exaggerated, as there have been no cases of wars being waged solely by killer robots, and the number of casualties inflicted by drones in recent conflicts, e.g. in

[68] Horowitz M. C. (2018), *Artificial Intelligence, international competition and the balance of power*, Texas National Security Review, Vol. 1, No. 3, pp. 37-57.
[69] Gilli A. e Gilli M., Artificial Intelligence and International Security, *working paper*.
[70] Ibid.
[71] Gilli A. and Gilli M. (2021), 'Technology and new global conflicts', *Aspenia*, Vol. 94, pp. 117-127.

Syria or Nagorno-Karabakh, are considerably lower than those caused by traditional clashes.

These fears and apprehensions, more generally, exacerbate the factual reality of the development of UAVs and artificial intelligence per se. In fact, both of the two meta-approaches used in artificial intelligence for autonomous weapons have limitations. The deductive, or top-down approach on the one hand, relying on *ex-ante* planning, which must include all kinds of eventualities or contingencies, is clearly unattainable with today's technology, especially if the killer robot has to act in a constantly changing environment. The inductive, or bottom-up approach, on the other hand, is based on collecting huge amounts of data in order to extrapolate trends and patterns that the software will have to learn through machine learning systems. The issues here are manifold: having access to these huge amounts of data, the high training costs to be incurred and finally the exposure to increased vulnerabilities, such as cyber-attacks, leading to a high exposure to operational risks and the consequent caution before their full use on the battlefield.[72] This second approach also has hardware limitations. The current computer architecture cannot handle the enormous amounts of data produced and does not have enough computational power to process them, given the recent explosion of deep learning as a new basis for AI development.[73]

In the light of the technological and structural limitations associated with the use of AI, excessive pessimism and apprehension about a technology that is still under development and far from disrupting the military field is evident.

[72] Gilli A. and Gilli M. (2021), 'Technology and new global conflicts', *Aspenia*, Vol. 94, pp. 117-127.
[73] Gilli A. e Gilli M., Artificial Intelligence and International Security, *working paper*.

4. ETICHAL AND LEGAL QUESTIONS

The acceleration in digitisation, the robotic era and autonomous systems, in which technology and robots seem to replace humans, and to be able to act or make predictions on their own, make the domain of ethics and the role of international regulations acquire renewed value and importance. There are usually two types of approaches to technological innovation: innovate first and then manage the consequences or try to prevent the risks of innovation *ex ante*. In many cases, the former approach has been used. However, as far as AI is concerned, the latter seems to find greater consensus given the debates on ethical principles.[74] In this last section, we will focus on these ethical and regulatory challenges or attempts at regulation.

The increasing focus on ethics would require such new technologies, whether for commercial or military use, to behave according to the norms, values and judgements that human beings would adopt, following what we might call our moral code. However, their behaviour and actions depend on the codes entered during programming, the algorithms and the data they possess. Despite increasing ethical and normative contributions by some groups, the domain of AI ethics lacks precise parameters that determine the relationship between technological development and broader societal discussions.[75] Consequently, academics, scholars, and NGOs, demand that the realisation of UAVs be carried out following certain principles and guidelines, which we can encapsulate in six assumptions.[76]

Foremost among these is that AI has human beings and human rights at the centre of its attention and programming, so that its use is aimed at improving, and not undermining, the living conditions of human beings. Secondly, in order to develop a sense of trust in the use of AI, their actions must be explainable, understandable and transparent, *ex ante* and *ex post*, whether they are lethal autonomous weapons or driverless cars. A third position is the problematic nature of the attribution of responsibility and accountability. This criticality arises especially in the case of errors, malfunctions or cyber attacks, which complicate the scenario even more, increasing the difficulty in identifying those responsible. The allocation of responsibility in the event of misunderstandings is linked to the additional principles of reliability and security, which must be ensured prior to the use of new technologies, all the more so in the case of stand-alone systems.

---

[74] Luca de Biase (2021), "Where ethics meet algorithms", *Aspenia*, Vol. 94, pp. 128-135.

[75] Daniel Zhang, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, James Manyika, Juan Carlos Niebles, Michael Sellitto, Yoav Shoham, Jack Clark, and Raymond Perrault, *The AI Index 2021 Annual Report*, AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, CA, March 2021.

[76] Gilli A. (2020), *"NATO-Mation": Strategies for leading in the age of Artificial Intelligence,* NDC Research Paper N. 15, Rome, NATO Defence College, pp. 29-34.

These principles, to be followed in the planning, construction, testing, evaluation and validation phase, before implementation, would not only reduce the margin of direct risk to human life, but also slow down the dissemination or commercialisation of such technologies. The fifth position in this debate envisages and calls for the integration of principles of equality and inclusion. As in any competition, depending on the capabilities of individual countries, there are companies or individuals with less digitisation, data accessibility or new technologies. The risk involved is to change the hierarchical order or to increase the differentiation and marginalisation of some groups of individuals, or countries, to the detriment of others. Finally, the last position concerns privacy and data governance issues. These principles are among the most controversial and debated because they are in opposition to the principles of security and validation, as the latter would require the collection of huge amounts of data, while privacy rules demand their reduction and stricter regulation.[77]

Some general rules, however, to define the legal requirements that a weapon system must meet and achieve before its implementation already exist. The most generic framework for the review of new weapons, prior to their entry into use, is Article 36 of the Additional Protocol (I) to the Geneva Convention 1977, according to which *"In the study, development, acquisition or adoption of a new weapon, new means or methods of warfare, a High Contracting Party shall be obliged to determine whether its use is not prohibited, in certain circumstances or under any circumstances, by the provisions of this Protocol (I) or by any other rule of international law applicable to that High Contracting Party"*.[78] The two substantive rules that must be a part of Article 36 and the review of new weapons are the rule against non-discriminatory weapons, i.e. the banning of all new weapons that cannot distinguish their targets, differentiating civilians from military following the principle of distinction, according to Article 54 (3b) of Protocol (I), and Article 35 (2) of Protocol (I) prohibiting any type of weapon whose nature may inflict unnecessary and superfluous violence or harm, following the legal principle of proportionality.[79] According to the views of the most sceptical, however, this international legal framework is not sufficiently regulatory for autonomous weapon systems.

At the centre of the regulatory and legal debate concerning autonomous systems, the main focus is not only on a more general regulation of AI, but also

---

[77] Gilli A., Pellegrino M. e Kelly R. (2019), "Intelligent Machines and the Growing importance of Ethics", in Gilli A. et al., *The Brain and the Processor: unpacking the challenges of human-machines interactions*, NDC Research Paper, Rome: NATO Defence College; Luca de Biase (2021), "Where ethics meet algorithms", *Aspenia*, Vol. 94, pp. 128-135.

[78] United Nations (1977), *Protocol Additional to the Geneva Conventions of the 12 August 1949, and relating to the protection of victims of international Armed Conflict (Protocol I),* of 8 June 1977, UN.

[79] Anderson K. e Waxman M. (2013), *Law and ethics for autonomous weapon systems, Why a ban won't work and how the laws of war can*, Task Force on National Security and Law.

finds the most sceptical and pessimistic positions towards Lethal Autonomous Weapon systems, or LAWs. There are generally four main criticisms that are repeated and used against LAWs to demand their ban. The former argues that programming systems will never reach the ethical, moral and legal standards required for their use on the battlefield. An *a priori* assumption based on levels of development that they will never be reached, according to the proponents of this position, thus excluding the possibility of progress and refinement that AI could develop in the long run. The second position, on the other hand, is against the possibility of partially or completely excluding the presence of human control, i.e. moral agents, in war scenarios and beyond. This criticism is broader and would require a collective, national and international decision on the desirable level of development of autonomous systems to be achieved. In fact, driverless cars, to give an example, have already entered the commercial space and are generally accepted. The third argument relates to one of the ethical principles that these new technologies should follow, namely the problem of attribution of responsibility. In the case of errors, malfunctions, targeting of civilians and thus war crimes, determining the person materially and legally responsible for such actions, would become even more complicated if not impossible. This problem may perhaps one day be solved by improving programming, deep learning or available data, but an excessive focus on limiting AI risks slowing down even those developments with potential positive results. Finally, the last objection argues that the acceleration in the development of autonomous weapon systems will lead to an increase in armed conflicts and confrontations in the long run, since the greater precision in weapon systems and the possibility of reducing the physical presence of soldiers, and thus collateral damage to them and civilians, would decrease the disincentive for wars.[80]

These ethical-regulatory examples present a scenario that is still being defined and therefore incomplete. The major European and Atlantic powers have an interest in making the development and use of these new technologies legally regulated. While on the one hand we find the more sceptical and pessimistic calling for a ban on killer robots, on the other hand there may be room for multilateral action that is more regulatory than restrictive, even though reaching an agreement, especially for new technologies of both commercial and military use, entails various difficulties.

Arguments in favour of banning AI focus more on the consequences it could have, particularly when applied to the military domain. However, the still premature stage of development of this technology does not justify this widespread pessimism and excessive limitations risk harming future progress.

[80] Anderson K. e Waxman M. (2013), *Law and ethics for autonomous weapon systems, Why a ban won't work and how the laws of war can*, Task Force on National Security and Law.

Conclusions:

- Most of the critics to the use of AI in the military domain are based on three assumptions which present clear limits.
- The widespread pessimism regarding this new type of technology is not supported nor legitimated by the current stage of AI's development.
- The development of the ethical and legal framework for the use of AI within the military domain can be further improved by multilateral action.

Chapter IV

NEW WARFARE: POTENTIAL RISKS AND MITIGATIONS

by Enrico Savio and Enrico Comin

SUMMARY: The development and application of disruptive technologies are already redefining Defence and Security scenarios. Scenarios in which - soon - speed, efficiency and precision of military operations will reach levels never imagined before, and where the increasingly sophisticated capabilities of the machine may exceed the human element, its control and its responsibility (man-out-the-loop). This is a radical change in tactical and strategic paradigms, with profound ethical and moral implications. This chapter aims first to illustrate and analyse the state of the art of these technologies and their potential impact in future warfare. At the same time, it aims at highlighting the risks, the main challenges and the possible strategies to manage the evolutions already underway, with the full awareness that technology must be oriented and developed within the framework of shared norms, values and ethical and moral principles.

1. DISRUPTIVE TECHNOLOGIES IN DEFENCE

Technological development has always driven profound changes in the characteristics of armed conflicts and geopolitical balances. The last few years, in particular, have witnessed substantial changes in the global technology ecosystem due to ever faster progress and an increasing rate of deployment. Innovation today, compared to the past, is in fact increasingly driven by the commercial sector and is contaminating - in an increasingly pervasive manner - the defence sector, giving rise to new security challenges never before conceived.

As outlined in the 2018 US *National Defence Strategy*, the security scenario of the future will be profoundly influenced by both rapid technological advances and the changing nature of warfare. New technologies, including advanced computing, ”big data” analysis, artificial intelligence, autonomous systems, robotics, directed energy, hypersonics, and biotechnology will be the critical junctures on which the wars of the future will be fought and won.[81] in a report for the *Center for a New American Security*, Ben FitzGerald and Shawn Brimley defined *Disruptive Technologies* in the defence sector as "a technology or set of technologies applied to a major problem in a way that radically alters the symmetry of military power among competitors" and that "immediately transcends the policies, doctrines, and organisation of all actors involved".[82]

[81] United States Department of Defence (2018), *“Summary of the National Defence Strategy”*.
[82] FitzGerald, B. Et al. (2013), *“Game Changers: Disruptive Technology and U.S. Defence Strategy”*, Center for a New American Security.

The concept of disruptive technology, however, is not a product of our times but has always been valid throughout history. In fact, as noted by the French historian Jacques Le Goff, from a strictly military point of view, the thesis of a cavalry born 'naturally' during the 8th century out of the need to counter the rapid incursions of the Arabs of Spain must be discarded. It was the invention of an object, the stirrup, which, by allowing greater stability in the saddle - and thus a more efficient attack - fundamentally revolutionised the role of the horseman both on the battlefield and in society.[83] Returning to our times, the permeation of these disruptive technologies into the global defence landscape is generating numerous and substantial questions regarding the management of future conflicts and, more generally, the balance of power between states. In this context, the aim of the following paragraphs is to list the main technologies considered to be disruptive in the defence sector with their definitions, providing an analysis of their main characteristics.

2. ARTIFICIAL INTELLIGENCE (AI)

Although there is no single, agreed definition of artificial intelligence, the term AI is generally used to refer to a computer system with human-level cognitive capabilities. AI is divided into two categories: Narrow AI and general AI. Systems falling into the former category can only perform the specific task for which they have been trained; whereas the latter, through autonomous learning, may one day be able to perform a wide range of tasks, including those for which they have not been specifically trained.

Narrow AI is currently embedded in numerous military applications, including but not limited to intelligence, surveillance and reconnaissance, logistics, computer operations, command and control and semi-autonomous and autonomous systems. These technologies are intended in part to support or replace human operators, who will be called upon to do more complex and cognitively demanding work. AI-enabled systems could react much faster than those relying on operator input, cope with an exponential increase in the amount of data available for analysis, and enable new operational concepts, such as swarming,[84] which could confer a tactical advantage by overwhelming adversary defensive apparatuses. Significant in this context is the result highlighted in

[83] Le Goff, J. (1997), *'L'Uomo Medievale',* Editori Laterza.
[84] Swarming refers to cooperative behaviour in which unmanned systems communicate, collaborate and coordinate autonomously by making collective decisions to achieve a specific task.

DARPA's research programme '*AlphaDogfight*', focusing on AI-powered air combat capabilities, where - in a series of simulated air duels between one aircraft piloted by artificial intelligence and another by a human pilot - the AI-controlled aircraft prevailed with astonishing results.[85]

AI, however, could introduce a number of cross-cutting challenges such as, for instance, vulnerability to cognitive biases and biases arising from the datasets on which the algorithms are trained.[86] Indeed, several researchers have repeatedly identified instances of racial bias in AI facial recognition programmes, mainly due to the lack of diversity in the images the systems were trained on, while some language processing programmes have developed gender biases.[87] This type of vulnerability could have significant implications for applications of AI in a military context. For example, unconsciously incorporating undetected biases in testing could lead to cases of misidentification of targets. In military systems, such algorithms could produce unpredictable and unconventional results capable of generating unexpected failures. Moreover, these vulnerabilities could be intentionally exploited by malicious actors or adversaries to disrupt the identification, selection and engagement of targets through or with the support of AI. This could, in turn, raise ethical concerns, or potentially lead to violations of armed conflict laws, if the system selects and engages a target or class of targets that has not been approved by a human operator.

Recent news and analyses have further highlighted the role of AI in enabling falsifications and digital manipulations of photos, audio and video with increasingly realistic results. Such fictitious products are known as deep fakes.[88] These AI capabilities could be used as part of operations aimed at undermining information capabilities. Indeed, deep-fake technology could be used to generate fake news, influence public opinion, erode public trust and attempt blackmail of government officials. For this reason, some analysts argue that social media platforms, in addition to employing deep fake detection tools, should strengthen content classification and authentication solutions.[89]

---

[85] Hitchens, T. (2020), *"AI Slays Top F-16 Pilot In DARPA Dogfight Simulation"*, Breaking Defence, https://breakingdefence.com/2020/08/ai-slays-top-f-16-pilot-in-darpa-dogfight-simulation/.

[86] Training data are data used to teach AI models or machine learning algorithms to make correct decisions. They can be supplemented by successive sets of data called validation and test sets.

[87] Congressional Research Service (2020), *"Emerging Military Technologies: Background and Issues for Congress"*.

[88] Deep fakes are a type of artificial intelligence used to create convincing and often indistinguishable images, audio and video. The term, which describes both the technology and the resulting fake content, is a portmanteau of the words *deep learning* and *fake*.

[89] Ivi; Roth, Y., Achuthan, A. (2020), *"Building rules in public: Our approach to synthetic & manipulated media"*, Twitter, https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media.html.

To further complicate the problems of predictability and security, the types of AI algorithms with the highest performance are currently unable to explain their processes. For instance, Google created a cat identification system that achieved impressive results in identifying felines on YouTube. However, none of the developers of the system were able to determine which traits of a cat the tool used in its identification process.[90] This lack of so-called 'explainability', which is common to most AI algorithms, led the Defence Advanced Research Projects Agency (DARPA) to make five-year research investments to produce "explainable" AI' tools. Insufficient explainability may create further problems in a military context, as the opacity in the functioning of the algorithm may lead operators to have over- or under-confidence in the system. In addition to this, it can call into question several other steps in human-computer interaction, including:

- goal alignment → man and machine must have a common understanding of the goal, which, in a dynamic environment, tends to change, necessitating simultaneous adaptation by both man and machine, based on a shared picture of the environment;
- task alignment → man and machine must understand the boundaries of each other's decision space, especially when the goals change. In this process, humans must be fully aware of the limitations of the machine design to avoid placing inappropriate confidence in the system;
- human-machine interface → due to the requirement for timely decisions in many military AI applications, traditional interfaces can slow down performance. It is therefore necessary to consider solutions to ensure real-time coordination between man and machine.

Finally, explainability could challenge the ability to verify and validate the performance of AI systems prior to operational use. In fact, the current lack of an explainable output does not allow for the generation of an audit trail for tests aimed at verifying performance standards. Increasing the ability to explain cognitive processes will therefore be one of the key activities to raise confidence in such systems to appropriate levels. In all cases, the security of Artificial Intelligence from a cyber perspective (in its design, training and operations) is of paramount importance to ensure that it behaves in accordance with how it was designed and trained. This priority is, among other things, a clear support for the need to have technological sovereignty in this field and, consequently, to devote adequate attention and funds to it.

[90] Congressional Research Service (2020), *"Emerging Military Technologies: Background and Issues for Congress".*

3. QUANTUM TECHNOLOGIES

The almost unimaginable increase in the rate and order of computation enabled by quantum technologies (see Chapter II) would provide profound advantages in strategically vital areas, including cryptography and decryption, radar and sensor technology, navigation and targeting, simulation and data mining, machine learning and pattern recognition.

In general, these technologies have not yet reached the necessary maturity for their actual use in the military, but they could have significant implications for the future of communications, cryptography and stealth technologies.[91]

In particular, so-called quantum communications could enable the development of secure transmissions that cannot be intercepted or deciphered. Quantum technology would be able to offer numerous other applications in defence, such as quantum radar systems that - it is assumed - will be able to identify performance characteristics (e.g. cross-sectional area and speed) of objects with a higher level of accuracy than conventional radar systems.[92] If realised, these systems would significantly facilitate the tracking and targeting of low-observable, or stealth, aircraft. Similarly, advances in quantum sensing could theoretically enable significant improvements in the detection of all those platforms operating below sea level, making the oceans "transparent".[93]

However, the military application of these technologies would be limited by the fragility of the quantum states, which can be disrupted by minimal movements, temperature changes or other environmental factors. As the physicist Mikkel Hueck explained, "If future devices using quantum technologies require cooling at very low temperatures, this will make them expensive, bulky and energy-hungry". Consequently, widespread adoption is likely to require significant advances in materials development and manufacturing techniques.[94]

---

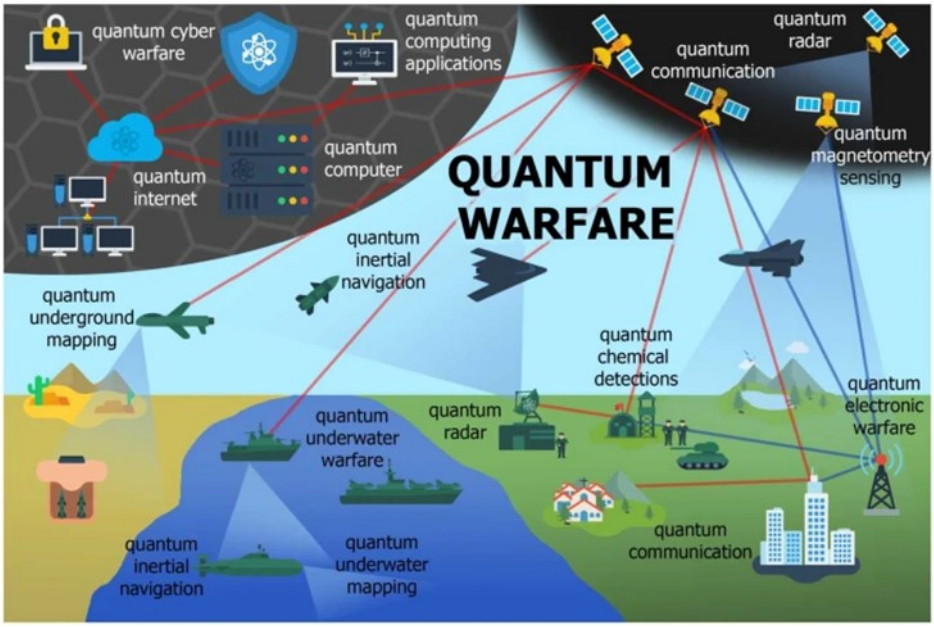
[91] Stealth technology, also known as low-observable technology, is a sub-discipline of military tactics and passive and active electronic countermeasures that covers a range of methods used to render personnel, aircraft, ships, submarines, missiles, satellites and ground vehicles less visible.

[92] Krelina, M. (2021), "*Quantum technology for military applications"*, EPJ Quantum Technology 8, 24, Springer Open.

[93] Congressional Research Service (2020), *"Emerging Military Technologies: Background and Issues for Congress".*

[94] Ibid.

Figure 1, Possible military applications of quantum technologies. Source: EPJ Quantum Technology.



Quantum technologies have the potential to profoundly influence many areas of human endeavour, particularly defence, where they will increase the sensitivity and efficiency of instruments, introduce new capabilities and refine modern warfare techniques. As illustrated in Figure 1, the possible applications of quantum technology for defence, security, space and intelligence in different aspects of new warfare are extremely numerous. However, it is important to consider that many applications are still more theoretical than realistic.

The significant quantum progress achieved in the laboratory does not always translate into similar progress outside it. Real deployment also involves other aspects, such as portability, sensitivity, resolution, speed, robustness, and cost; not to mention that integrating quantum technology into a military platform is even more challenging, given the additional requirements it would have to fulfil. Apart from quantum computers - most of which will be located in data centres in a way similar to those for civil use - the integration and implementation of quantum sensing, imaging and networks face several challenges, posed by the increased

demands of military use (compared to civil, industrial or scientific requirements).[95] Moreover, this field is still in an embryonic state. Further discoveries, whether positive or negative, could generate further advantages or disadvantages.

4. LETHAL AUTONOMOUS WEAPONS (LAWS)

Although there is no internationally agreed definition of lethal autonomous weapon systems, the US Department of Defence defines LAWS as a class of weapon systems capable of both autonomously identifying a target and employing a weapon to engage and neutralise it, without human control.[96] Autonomous weapon systems, therefore, are capable of completing specific tasks autonomously, without operator input.

Although these systems are not yet seeing widespread development, they are expected to play a key role in operating environments in which traditional systems may not be able to operate. This level of autonomy is also known as *man out of the loop* or "full autonomy". Such systems, which are defined on the basis of their level of operational autonomy, can be either human supervised, or man-on-the-loop, in which operators have the ability to monitor and stop the weapon's engagement; or be semi-autonomous systems, referred to as man-in-the-loop, which only engage individual targets or specific groups of targets, which have been selected by a human operator.

LAWS use AI computer algorithms and sensor suites to classify an object as hostile, make an engagement decision and deploy a weapon on the target. As emphasised by several analysts, autonomous weapon systems could allow military targets to be hit more accurately, thus reducing the risk of collateral damage or civilian casualties.[97] Some 25 countries and 100 non-governmental organisations have called for a preventive ban on LAWS, motivated by ethical concerns, such as a possible perceived lack of accountability for use and a possible failure to comply with the requirements of proportionality and distinction.[98] Some analysts have also raised concerns about the potential

---

[95] Van Amerongen, M. (2021), *"Quantum technologies in defence & security"*, NATO Review, https://www.nato.int/docu/review/articles/2021/06/03/quantum-technologies-in-defence-security/index.html.

[96] United States Department of Defence Directive 3000.09 (2012-2017), *"Autonomy in Weapon Systems"*.

[97] U.S. Government (2018), *"Humanitarian Benefits of Emerging Technologies in the Area of Lethal Autonomous Weapons"*, United Nations Convention on Certain Conventional Weapons.

[98] Congressional Research Service (2021), *"Defence Primer: U.S. Policy on Lethal Autonomous Weapon Systems"*.

operational risks posed by lethal autonomous weapons, which take the form of "hacking, manipulation of behaviour by the enemy, unexpected interactions with the operational environment, or simple software malfunctions and errors".[99] Such risks could be present in automated systems, and be intensified in autonomous systems, in which if the human operator is unable to physically intervene, undesirable consequences, such as wider destructive effects, instances of friendly fire and civilian casualties, could be generated.

Another fundamental dimension of autonomous systems is their degree of complexity - both that of the system itself and that of the environment in which it operates - which influences the human operator's ability to predict and control its behaviour. In general, simpler systems - operating in simpler environments - will be easier to predict and, although probably more limited in the types of operations they can perform, their operation will presumably be more transparent to operators. However, the range of environments and situations in which they can operate is also likely to be more limited. To operate in a wide range of scenarios and perform more difficult missions, more sophisticated autonomous systems are required, which, by necessity, are inevitably more complex. This complexity can make the system less transparent in its processes even for trained operators. Consequently, predicting system behaviour, particularly when operating in the real world in complex, unstructured environments, can be more difficult.[100]

To limit such risks, autonomous systems should be tested and evaluated to ensure that they function as intended in realistic operating environments and against changing adversaries. They should also complete engagements in a timeframe consistent with the intentions of the commander and operator and, if they are unable to do so, terminate their activities or seek further input from the human operator before continuing. Finally, such systems should be resilient to the point of being able to minimise failures capable of causing unintentional engagement or loss of control, to the benefit of unauthorised actors. Any change in the operational state, e.g. due to machine learning, would require the system to go through a testing and evaluation process again to ensure that it has retained its security features and ability to function as originally intended.[101]

As technology advances, the risks of using such systems must be carefully considered. Much of the current debate focuses on legal, moral or ethical issues. However, autonomous weapons also raise important issues of controllability and

---

[99] Scharre, P. (2016), *"Autonomous Weapons and Operational Risk"*, Center for a New American Security.
[100] *Ibid.*
[101] Congressional Research Service (2021), *"Defence Primer: U.S. Policy on Lethal Autonomous Weapon Systems"*.

safety, especially in the event of failure. Over a long enough period of operational deployment, some failures are inevitable and using autonomous weapons would mean accepting the consequences of these failures[102].

There is a need for greater transparency among states on how to deal with autonomy in weapon systems. Few countries have clear national policies on the use of such instruments. Given the potential for dangerous interactions between autonomous systems and the risks mentioned above, it is particularly urgent to arrive at an internationally consistent and shared regulation on their use. The confrontation and competition between the Great Powers in terms of Defence and Security naturally leads to a race towards greater efficiency of their Armed Forces, an enhancement that passes through greater operational speed, which, in turn, requires greater automation, pushing this race to a further acceleration of the pace of battle. The result of this continuous increase in speed and automation could lead to an unstable situation, when unexpected interactions between autonomous systems or hacking could lead to a 'blitzkrieg', causing any conflict to quickly spiral out of human control.

5. HYPERSONIC

Hypersonic weapons are defined as all those weapons that are capable of moving at a speed of at least Mach 5, which is five times the speed of sound. Most conventional ballistic missiles fly at hypersonic speeds, while conventional cruise missiles generally fly at subsonic (less than Mach 1) and supersonic (Mach 1 to 5) speeds. In practice, the term 'hypersonic weapons' refers to weapons that fly at a lower altitude than intercontinental ballistic missiles, higher than traditional cruise missiles, and are largely intended for regional rather than intercontinental use.[103] There are two categories of hypersonic weapons: hypersonic glide vehicles (HGVs), which are launched by a rocket before gliding towards a target; and hypersonic cruise missiles (HCMs), which are powered by high-performance engines for the duration of the flight.[104] Hypersonic glide vehicle (HGV) systems are typically launched with a rocket into the atmosphere and released at an altitude between 40 and 100 km from where they glide towards their target. HGVs have a range comparable to ballistic missiles but fly at a lower altitude. A negligible portion of their flight path follows a ballistic trajectory. In addition, such systems are manoeuvrable during the glide phase and can be redirected in flight

[102] Scharre, P. (2016), *"Autonomous Weapons and Operational Risk"*, Center for a New American Security.
[103] Congressional Research Service (2020), *"Emerging Military Technologies: Background and Issues for Congress".*
[104] Ibid.

to a target other than the one initially planned. Hypersonic cruise missiles (HCM), on the other hand, as they are powered for the entire flight, must be propelled to a speed of around Mach 5 before a jet engine (ramjet, scramjet) can take over to maintain it. HCMs could be launched from the ground, from the air or from a ship and would probably fly at an altitude of 20-30 km, beyond the range of most current air-to-surface missile defence systems and would be able to reach targets 1,000 km away within minutes.[105]

Ballistic missiles follow a largely predictable parabolic ballistic trajectory, flying high above the atmosphere before plummeting back to Earth. This allows those at the receiving end to more easily track the ballistic missile - in its intermediate phase of flight - through radar and derive reasonable predictions of where the warhead will land. HGVs, on the other hand, do not follow a parabolic ballistic trajectory and can be manoeuvred en route to the target, significantly decreasing the predictability of the trajectory and thus making it difficult to apply countermeasures and possible defence. Moreover, they are supposedly less detectable by radar due to their low altitude route as compared to that of ballistic missiles.

There are conflicting views on the strategic implications of hypersonic weapons. Some analysts have identified two factors that could have significant consequences: 1) the weapon's short flight time, which in turn compresses response times, and 2) its unpredictable flight path, which could generate uncertainty about the weapon's target and thus increase the risk of miscalculation or inadvertent escalation in the event of a conflict.[106] Other analysts argue, however, that the strategic implications of hypersonic weapons are minimal. It would not change the already existing capability to strike with intercontinental ballistic missiles that, if launched en masse, could overwhelm missile defences.[107]

The development and potential future deployment of hypersonic weapons highlight a number of broader issues and questions that deserve attention. The implementation and adoption of such tools is fostering the research and development of technologies to defend against them. These include kinetic interceptors, electromagnetic railguns and high-powered lasers, which, as will be discussed in more detail in the next section, may also have potential use as offensive weapons. As mentioned earlier, hypersonic weapons may reduce response times in certain scenarios compared to current cruise and ballistic missiles, significantly compressing decision-making timeframes and thus

---

[105] Congressional Research Service (2021), *"Hypersonic Weapons: Background and Issues for Congress".*

Bugos, S. et al. (2021), *"Understanding Hypersonic Weapons: Managing the Allure and the Risks"*, Arms Control Association Report.

[106] Brehm, M., de Courcy Wheeler, A. (2019), *"Hypersonic Weapons – Discussion Paper for the Convention on Certain Conventional Weapons (CCW)"*, Article36.

[107] Ibid.

contributing to the trend of increasing reliance on artificial intelligence - both to inform human decision-makers and to automate certain processes - raising concerns about the risks inherent in time-pressured decision-making.

6. DIRECT ENERGY WEAPONS (DEW)

The US Department of Defence defines directed energy weapons (DEW) as those that use concentrated electromagnetic, rather than kinetic, energy to "disable, damage, or destroy enemy equipment, facilities, and personnel".[108] A DEW, therefore, is a weapon system that uses a targeting system to control the emission of energy (electromagnetic, laser, microwave, photonic energy and nuclear radiation) as a means of damaging or destroying equipment and structures or injuring enemy personnel. There are different designs for a DEW system, capable of producing different forms and levels of energy. These design variations are each more efficient according to their multiple applications: from low-energy systems for crowd control and temporarily interrupting the operation of electro-optical sensors, to high-energy systems for causing material damage. Some of these designs have already been introduced into service by various military forces around the world.

A DEW is not a detonation weapon. There is no explosion, kinetic energy, fragmentation, impact or penetration. There is only thermal energy and light, which cause effects and possibly non-kinetic damage. Explosion, fragmentation and additional thermal damage may occur as secondary effects resulting from the initial interaction between the energy and the target materials themselves. The thermal energy damage mechanism is based on the thermal properties of the materials of which the target is composed. If such materials are vulnerable to heat absorption - concentrated in a small area and over a short period of time - they are heated to the point of combustion or melting.[109]

These weapons fall into further categories:

- high-energy laser (HEL)
- high-power radio frequency (HPRF)
- high power microwave (HPM)

[108] United States Department of Defence (2020), *"Joint Electromagnetic Spectrum Operations"*, Joint Chiefs of Staff Publication 3-85.

[109] Spencer, M. (2020), *"Directed Energy Weapons: Playing with Quantum Fire"*, Australia Air Power Development Center.

Such weapons could be used by ground forces in short-range air defence (SHORAD), against unmanned aerial systems (C-UAS), or in missions against rockets, artillery and mortars (C-RAM). Direct energy weapons could also reduce ammunition costs and, assuming sufficient power supply is available, provide almost unlimited range.[110] In contrast to existing conventional systems, they could provide an effective means of missile defence or towards possible swarms of unmanned systems. Theoretically, laser-based DE weapons could also provide options for intercepting missiles being launched, given their travel time equal to the speed of light. However, as in the case of hypersonic missile defence, experts disagree on the accessibility, technological feasibility and usefulness of this application. High-power microwave weapons - a subset of directed energy weapons - could be used as non-kinetic means to disable electronics, communication systems and improvised explosive devices, or as non-lethal systems to neutralise targets.

Directed energy weapons have long captured military attention - and budgets - and are now on the cusp of technological maturity. While doubts remain as to whether some models can be fully operational, recent tests of DEW prototypes have shown that this form of weaponry has gone beyond the theoretical concept. As the underlying technology matures and undergoes testing outside of laboratories, it is likely to attract more attention from armed forces and governments that aspire to achieve technical superiority over their adversaries, including developing solutions for eventual deployment in space.

7. SELF-DRIVING SYSTEMS

Advances in disruptive technologies influence the field of robotics and autonomous capabilities, changing the military balance and the dynamics of warfare now and in the future. Unmanned systems, with a varying degree of autonomy, are now a common factor for the world's armies and, in their air, sea and land forms, are used to perform various types of missions: intelligence, surveillance and reconnaissance (ISR), search and rescue (S&R), logistics, mine clearance and destruction of improvised explosive devices (IED), armed patrols and targeted target destruction.

A current example of the deployment of such systems is the conflict in Ukraine, where the Kiev army is countering Moscow's troops through the use of

[110] Congressional Research Service (2021), *"Department of Defence Directed Energy Weapons: Background and Issues for Congress".*
Congressional Research Service (2020), *"Emerging Military Technologies: Background and Issues for Congress".*

Turkish “Bayraktar TB2” javelins and drones already deployed in 2021 by Azerbaijan against the Armenian army. On the other hand, the Russian army seems to be using the KUB-BLA, a “loitering munition” from ZALA Aero, equipped with artificial intelligence, which makes it capable of detecting, recognising objects and thus identifying targets in real time. The KUB-BLA, fired from a portable launcher, is able to fly for 30 minutes, reaching a speed of 130 km/h, and crash into an independently identified target.

The United States, China, Russia, the United Kingdom, Israel and Turkey have progressively increased their investments in this field in recent years, designing technologically advanced systems equipped with artificial intelligence and capable of striking targets hundreds of kilometres away. Such systems include fleets of ships, land vehicles, missiles, aircraft and drones guided by artificial intelligence. A well-known example of the use of such systems is that of March 2021, when a Turkish Kargu-2 drone was used in Libya to engage in autonomous attacks against ground targets such as soldiers and vehicles.

Advances in technologies such as artificial intelligence, robotics and data fusion are capable of revolutionising the use of self-driving systems, enabling large numbers of drones to operate in a collaborative, coordinated and responsive manner to achieve common goals. Although it is in the early stages of development and experimental application, the swarm is a concept that - if fully developed - can have profound tactical and strategic effects, offering greater defensive and offensive options to military forces.

Coordination and responsiveness are the distinguishing features of the swarm. Its component drones are interconnected and in constant communication with each other for real-time sharing of information from their sensors and collective decision-making guided by artificial intelligence. Each element of the swarm has a specific role in a larger system, which self-coordinates the actions of its elements dynamically. For instance, some elements may deploy their sensors for target location and tracking, others perform jamming and electronic warfare tasks, and still others engage hostile forces. As a whole, the swarm reacts dynamically to changes in the battle space, performing complex non-linear and counter-intuitive manoeuvres.

The effectiveness of the swarm, as mentioned, depends on the stable connection between its components and the proper functioning of the AI that governs it, thus making it vulnerable to spoofing, jamming, cyber-attacks or technical malfunctions. New efforts will therefore have to be made to develop effective countermeasures as the swarms are deployed and become more advanced. Given the inherent decentralisation of swarms and their ability to react quickly in complex ways, it cannot be ruled out that the counter-swarm will become another swarm.

Given the technical complexity and high cost of the necessary know-how, it is safe to assume that the swarm will be the prerogative of the major military powers, which will enjoy a considerable advantage both against regular forces lacking similar capabilities and against insurgents due to the swarm's ability to provide near-permanent and reactive monitoring over a large area. Finally, in light of the interest that major military powers are expressing in the potential of swarms of unmanned systems, it can be expected that this emerging technology will become a game-changer in the warfare of the coming decades.

8. THE IMPLICATIONS AND IMPACTS OF THESE TECHNOLOGIES ON CURRENT TACTICAL AND STRATEGIC PARADIGMS: THE TRANSITION TO HYPERWAR

The implications of the development and application of disruptive technologies in defence are manifold and are changing the “rules of the game”. The issues of speed, efficiency and accuracy, as well as the need for human control over AI-powered military capabilities and, last but not least, the ethical implications of using such tools represent the heart of the matter in the current debate on artificial intelligence applications and national security. Our increasingly interconnected societies are promoting a “democratisation” of technology: the rate of deployment and availability of high-tech solutions will be increasing and “multi-sectoral”. The application of these technologies to new tools and methods of warfare is stimulating an international arms race dynamic, both in conventional weapons and disruptive technologies, challenging current strategic paradigms and balances of power. Concepts such as the “left-of-launch” missile defence[111] and the disabling of nuclear command and control facilities by cyber means could reinforce a ”use it or lose it” approach[112] for related first strike capabilities[113].

Furthermore, several experts have raised concerns about the possibility of using deepfakes to manipulate and influence command and control processes

[111] The “left of launch” strategy is based on a pre-emptive strike with new non-kinetic technologies, such as electromagnetic propagation, cyber as well as offensive force to defeat nuclear ballistic missile threats before they are launched.

[112] *”*”Use it or lose it” is the idea that if a state were to fear that its nuclear arsenal could be neutralised, it would prefer to use its nuclear weapons at the start of a crisis before losing them.

[113] A “first strike”, also known as a preventive nuclear attack, is an attack aimed at an opponent's nuclear arsenal that effectively prevents retaliation against the attacker.

Vienna Center for Disarmament and Non-Proliferation (2021), *“Deterrence Stability and the 21st Century Technology Boom”*.

and early warning systems, as well as to alter the data on which the AI operates, causing unintended and potentially harmful consequences.[114] To respond effectively to the risks generated by such contingencies, the digital age requires the creation of high "cyber deterrence" capabilities. Indeed, combat is not only moving towards robotics, it is also becoming "ethereal". For instance, during its incursion into Georgia in 2008, Russia became the first nation to deploy cyberattacks on the enemy's command, control and communication systems to support a ground invasion. Similarly, to delay Iran's nuclear programme, the US and Israel allegedly launched the *Stuxnet* virus, which compromised the capabilities of centrifuges engaged in the uranium enrichment process. China has gained possession of large databases of information on US government personnel, as well as penetrating the networks of US defence contractors, airlines and technology companies.

These examples, which represent only "'the tip of the iceberg", illustrate the substantial advances in weapons technology over the past two decades, which observers sometimes refer to as a "revolution in military affairs".[115] With the advent of AI and other emerging technologies, however, the traditional definitions we are used to are set to change. At a fundamental level, battle, war and conflict are competitive processes in terms of time. Since time immemorial, human beings have sought to be faster in the competition resulting from combat, both in an absolute and relative sense. And in this regard, AI will drastically change the speed of warfare, not only altering the role of humans in conflict, but also leveraging technology like never before. And this is happening not only because technology is changing, but also because of the acceleration of its rate of transformation.

This is the central issue facing the preparedness for future armed conflicts. The actor who will be able to create, master and exploit a balance, quoting Clausewitz, between the nature of war and the character of war - especially within the new environment of AI, data analysis and supercomputing - will inevitably prevail in the conflict.[116] In a geopolitical scenario increasingly defined by new and emerging technologies, national defence is one of the most logical areas of development for the 21st century. It is therefore of paramount importance to assess the revolutionary impacts of artificial intelligence and other emerging technologies on every aspect of national security and armed conflict, including the accelerated pace of warfare and the critical role of continuous

---

[114] Singh Tanwar, S. (2020), *"Disruptive Technologies: Impact on Warfare & Their Future in Conflicts Of 21st Century",* Centre for Land Warfare Studies.

[115] Rabkin, J., Yoo, J. (2017), *"Disruptive Technologies to Upend Rules of War"*, National Defence Magazine.

[116] Allen, J., West, D. (2020), *"Op-ed: Hyperwar is coming. America needs to bring AI into the fight to win — with caution"*, CNBC. https://www.cnbc.com/2020/07/12/why-america-needs-to-bring-ai-into-the-upcoming-hyperwar-to-win.html

human control. With the advent of big data and deep learning, the third military revolution will enable a "digital synergy" of AI and other revolutionary technologies to collect, process and interpret zettabytes[117] of data in seconds. In the decision-centred conflicts of the future - where an adaptive network will form the connected "backbone" of commanders, operators and weapons - whoever is able to leverage the technological infrastructure to make the best decisions first will prevail.[118]

If this is the ultimate cognitive frontier, what is the human role in the mosaic of decisions made in seconds and battles waged in minutes? In their 2017 treatise "On Hyperwar", USMC General John Allen and SparkCognition founder Amir Husain redefine hyperwar as "a type of conflict in which human decision-making is almost entirely absent from the 'observe-orient-decide-act' (OODA) cycle.[119] As a result, the time associated with an OODA cycle will be reduced to almost instantaneous responses".[120] Hyperwar propagates the traditional OODA loop and expands monolithic decision-making structures into resilient, scalable and adaptable networks with the automated discretion to select, target and engage with adversary forces faster than their human counterparts.[121] Hyperwar is based on a system of systems and incorporates various levels of mechanised autonomy to release logistical constraints, coordinate movements, analyse data and recognise patterns. Machines will think and act faster than humans, adapting to the compression of operational time, in a context where speed and complexity will be key elements of the battlefield. A calculated and modular machine-enabled "decision stack" frees up command constraints to allow the human component to solve higher-level abstract problems. If exploited correctly, the ultimate benefit of mechanised military decision-making is an increase in "strategic bandwidth".[122] Indeed, while tactical decisions will be automated, strategic decisions will remain the domain of humans.

---

[117] The *zettabyte* is a unit of measurement of information or quantity of data, the term is derived from the union of the prefix zetta with byte and has the symbol ZB. "*Zetta"* is a so-called decimal prefix, the meaning of which is defined in the International System of Units, and corresponds to 10^{21} bytes or 1,000,000,000,000,000 (one trillion) bytes.

[118] Steel, C. (2021), *"Hyperwar: How Militarized AI is Transforming the Decision-Making Loop"*, American Security Project.

[119] The OODA cycle is the observe-orient-decide-act cycle, developed by military strategist and US Air Force Colonel John Boyd. Boyd applied the concept to the process of combat operations, often at the operational level during military campaigns. The approach explains how agility can outweigh power in dealing with human opponents. The concept is also particularly applicable to cyber security and cyber warfare.

[120] Allen, J. R., Husain, A. (2017), *"On Hyperwar"*, U.S. Naval Institute.

[121] Ibid.

[122] Steel, C. (2021), *"Hyperwar: How Militarized AI is Transforming the Decision-Making Loop"*, American Security Project.

Hyperwar, or combat conducted under the influence and support of AI - where human decision-making is almost entirely absent from the observe-orient-decide-act (OODA) cycle - is already beginning to permeate military operations. The human dimension of warfare will be put to the test in such a future hyper-warfare ecosystem, so considerable effort will be required in recruiting, educating, training and mentoring human talent.

The understanding of what decisions and choices are needed to build and operate in the new operational scenario - which includes these types of systems and requires these types of implementation decisions - can be found embryonically in some levels of the defence community. However, this decision-making capability will have to be built within a system that includes both Defence and Industry - in a mutually understandable and agreed way - in order to avoid the rejection of absolutely necessary solutions, performance doubts that could jeopardise operations or, worse, gaps that could occur where applications do not include the features that one party believes have been realised or executed by the other party. In fact, while industry is relatively quickly exploiting the “digital revolution”, starting with processes and operations, digital transformation has not yet been fully addressed in Defence, in all its facets and issues, including ethical ones.

9. DETERRENCE IN THE FUTURE HYPERWAR CONTEXT

While much of the debate on weapons enabled by disruptive technologies has focused ontheir humanitarian impact, such tools also raise considerable questions of strategic stability and deterrence. States have a long history of cooperating to regulate, prohibit or develop common standards and expectations for a variety of systems with a high impact on international stability. However, this convergence of advanced technologies could significantly increase a nation's military capabilities and thus generate profound changes in the international strategic balance, changing the preconditions for escalation and making the possibility of a “blitzkrieg” plausible.

From the crossbow to nuclear weapons, the development of military technology has always aggravated the security dilemma.[123] Of any progress in the development of war technology brings with it uncertainty as to how it will be deployed or how powerful it will be, new disruptive technologies introduce an even greater degree of uncertainty as to what capabilities they will be able to generate. For example, at the beginning of the Cold War, both the US and the Soviet Union were aware of the destructive capabilities of nuclear weapons and feared that the adversary might develop more powerful ones. The result was the nuclear arms race. AI, for its part, generates both forms of uncertainty: no one yet knows exactly how the AI-enabled weapons will be used on the battlefield, let alone how powerful they will be.[124] Artificial intelligence also introduces an additional layer of uncertainty by virtue of being an enabling technology. In fact, rather than constituting a single weapon system in itself, AI can be incorporated into multiple systems and infrastructures such as command and control centres and logistics solutions. Yet, it is not easy to determine how these innovations will change the nature of war conflicts. What effect will swarms of unmanned submarines have on naval warfare? What will happen when AI is not only integrated into existing armaments and command and control centres, but embedded in them through a bottom-up process? Which army will be able to integrate AI most efficiently and quickly into its weapon systems and tactics, and with what advantage on the battlefield? This uncertainty about how AI will be deployed and how effective it will be thus generates significant challenges for states, particularly in military strategy.[125]

[123] Meserole, C., (2018), *“Artificial intelligence and the security dilemma”*, Brookings.
[124] Ibid.
[125] Ibid.

Emerging and disruptive technologies are challenging the way deterrence, defence and, more generally, security strategies are formulated and applied at national and multilateral levels. Territorial size is no longer the main factor in determining the power of a state. Technological development, agility of manoeuvre and the speed and accuracy of decision-making will count for more than the available resources.[126] Therefore, a small player - but with considerable technological capabilities - might be able to successfully challenge a major power. New technologies will not only determine the way future conflicts will be fought and won, but will also have disruptive effects on the strategy of the great powers and on warfare itself, inevitably placing the emphasis on the process of revolutionising military technology. In this regard, in addition to the effect of such technologies in the warfare sphere, special attention must also be paid to the implications arising from those systems that are dual-use and the prerogative of entities and innovations in the commercial sphere.

Moreover, from the perspective of strategic stability, the assessments undertaken by states to determine a minimum credible deterrent may also change. These estimates will change significantly in the hyperwar era, as will the types and degrees of protection that will need to be provided to strategic systems and infrastructures.

The situation is made even more complicated by the multipolar environment of competition and conflict. An action intended to discourage an adversary might produce unexpected concerns among the other actors.[127] Nuclear deterrence has produced strategic stability through a combination of negotiation, public declarations and programmes aimed at arms acquisition. In the current and near future environment, acquiring more weapons will not produce more stability and the ability to negotiate on strategic and arms control issues with adversaries is significantly reduced compared to the past. Finding a way to coordinate this new strategic environment and strengthen international stability is not intuitive. In fact, as states are seeking to expand deterrence against emerging risks and new non-nuclear weapons, the old paradigm of stability is now compromised and will have to be renewed taking into account weapons enabled by disruptive technologies and their effects on deterrence.

---

[126] Hasan, S. (2020), *"How Disruptive Technologies Affect Deterrence, Defence and Security"*, Beyond the Horizon.

[127] Lewis, J. (2014), *"Disruptive Technologies and the Future of Deterrence"*, James Lewis, Center for Strategic and International Studies.

Dandashly, A. Et al. (2021), *"Multipolarity and EU Foreign and Security Policy: Divergent Approaches to Conflict and Crisis Response"*, JOINT Research Papers No.6.

As pointed out by several observers, there is a tendency to think of the conflict in a linear way as a direct path towards the escalation of the crisis.[128] However, technological complexity may allow the actors in the field to skip some steps in the process. Indeed, there could be a departure from predictable paths and a move towards a "wormhole" dynamic of crisis escalation, in which competing states could move between sub-conventional and strategic levels of conflict.[129] The idea of non-linear escalation paths gains further plausibility when non-state actors are involved and attribution becomes a long and complex process, requiring a mix of technical, social and political expertise.[130] In light of this complexity, the factors that encourage *hyperwar* may lend themselves to the development of a kind of "hyper-coercion", in which the ability to predict and anticipate an opponent's moves will be plausible.[131] In providing political and military decision-makers with alternative tactical and strategic options - based on a wide-ranging assessment of an unimaginable amount of data and information - AI can persuade the actors in command to delegate certain tasks to machines, necessitating a reconsideration of current assumptions and plans regarding automation in warfare.

Furthermore, AI can generate - in the long run - situational awareness based on predictive analysis.[132] In fact, through the fusion and analysis of information on the past and current behaviour of opponents, the AI will enable a concrete ability to anticipate their possible moves. Defenders will thus be able to respond pre-emptively by influencing and discouraging posture and possibly deterring the opponent from pursuing certain hostile actions.

AI thus introduces the ability to influence military deterrence and coercion in unique ways. It can alter cost-benefit calculations by eliminating the fog of war, imposing rationality on policy decisions, and decreasing the human cost of military engagement.[133] It can recalibrate the balance between offensive and defensive measures, tipping the balance in favour of prevention, and undermine existing assumptions about conventional and nuclear deterrence. In other words, AI could provide users with the ability to act on information collected, synthesised and processed in real time, increasing the certainty and severity of coercive

---

[128] Kubiak, K., Mishra, S. (2021), "*Emerging & disruptive technologies and nuclear weapons decision making: Risks, challenges & mitigation strategies*", European Leadership Network.
[129] Ibid.
[130] Ibid.
[131] Wilner, A., Casey, B. (2020), *"New Technologies and Deterrence: Artificial Intelligence and Adversarial Behaviour"* in *"Deterrence in the 21st Century-Insights from Theory and Practice"*, Netherlands Annual Review of Military Studies.
[132] Ibid.
[133] Ibid.

strategies and compressing the distance between intelligence, policy decisions and coercive action.[134]

Generally speaking, the context of the near future therefore requires the maintenance of a credible technological primacy, capable of fuelling an effective deterrence that induces possible aggressors to carry out - before their respective initiatives - a cost-benefit assessment. Without this primacy, the existing asymmetry between autocratic and democratic systems will hardly be mitigated by any diplomacy and form of international law. The other asymmetry to be considered, on a strategic level, is the imbalance - for a long time - between countries that are already technologically advanced and those that are still lagging behind in the race for technological advancement. Far from being able to control and counter more primitive realities, hyperwar risks being confronted with the most analogical and basic human beings can use in conflict. This is why technological awareness must be accompanied by an initial understanding of geo-economic and geo-political ecosystems. Without an implantation of awareness, it will be difficult to maintain a credible threshold of deterrence in defence of the value system to which the different communities of the globe feel they belong, especially if they are hinged on a democratic and peaceful structure. Unlike nuclear deterrence, which today can be described as rather static in its destructive potential, technological and digital deterrence will find constant evolution and consequent adaptation experiences from the scientific and military communities.

[134] Ibid.

10. DISRUPTIVE TECHNOLOGIES, DEFENCE AND SECURITY: ETHICAL AND MORAL ISSUES

"Today, the key question for humanity is whether to start a global AI-enabled arms race or prevent it from starting"[135]: this was the message at the heart of the letter signed by more than a thousand high-profile experts and researchers in the field of artificial intelligence - including Elon Musk (CEO and CTO of the aerospace company SpaceX, CEO and product architect of the automotive company Tesla), Steve Wozniak (co-founder of Apple), Demis Hassabis (CEO and co-founder of Google DeepMind), Stephen Hawking (cosmologist, physicist, mathematician and astrophysicist who passed away in March 2018) and Stuart J. Russell (professor of computer science at the University of California, Berkeley) - at the International Joint Conference on Artificial Intelligence, held in Buenos Aires in July 2015. *"AI has reached a point where the deployment of autonomous weapons is - practically if not legally - feasible within years, not decades, and the stakes are high: autonomous weapons have been described as the third revolution in warfare, after gunpowder and nuclear weapons"*[136]. AI, the document's subscribers argued, can be used to make the battlefield a safer place for military personnel. However, lowering the risk threshold for operators in the field could result in an incentive for offensive actions, leading to greater loss of life on the opposing side. In such a scenario, if a military power started to develop systems capable of selecting targets and operating without direct human control, it would start an arms race similar to the nuclear arms race.

This is a very clear-cut position that, however, in the years that followed, saw some of the letter's authors themselves change their orientation wholly or in part, but nevertheless unanimously recognise the enormous positive potential of Artificial Intelligence for the sustainable progress of the Planet.

Seven years after the Buenos Aires Conference, however, the central theme of their appeal remains open: will AI remove humans from decision-making processes altogether? This is not the case at the moment, as General John R. Allen, president of the Brookings Institution and former commander of NATO's International Security Assistance Force and U.S. Forces - Afghanistan (USFOR-A), explains. *“Despite its extensive use of drone technology, the United States today requires a briefed human being for every system it deploys. It is a moral point of great importance supported by American values and international norms,*

[135] Berkeley Engineering (2015), *“Open Letter on AI”*.
[136] Ibid.

and an intended limitation on the use of certain types of technology"[137]. The very fact of having to emphasise the centrality of the human element in the new scenario, which is increasingly characterised by the use of disruptive technologies, makes the impact that AI is having on US military doctrine all the more evident. In the past, the absolute symbol of US military superiority was the *"Strategic Triad'*, i.e. the three pillars of its nuclear deterrence strategy: missile squadrons, bomber fleets, and submarines equipped with ballistic missiles. Today, this triad has doubled, and to the three tools mentioned above we must add Artificial Intelligence, big data analysis and super computing. It is an all-round revolution, increasingly pervasive, that could - progressively and once and for all - replace humans *"from the process of environmental analysis, leading to more accurate, broader assessments and much faster reaction times"*[138]. In other words, the machine surpasses the human element, with the risk of bringing with it, in this quantum leap, well-defined responsibilities and choices based on ethical and moral principles, i.e. everything that makes our civilisation a long process of historical, cultural and social evolution.

On the other hand, awareness of technological development can become one of the keys to interpreting digital democracy. The passive approach of mere utilisation of the technological potential, in the confrontation between organised groups (but also between individuals, i.e. digital citizens), risks reproducing "the eternal chase between the shield and the spear", with the difference that human ingenuity and the ability to control it, direct it, and turn it to achievements risks being supplanted by the growing generating capacity of the machine. Together with awareness, therefore, the focus on the emotional factors of intelligence appears to be another moral pivot on which to graft part of the autonomous digital knowledge represented by AI. This is in order to maintain the primacy of the primal originality of the human being over the exponential current and empirical processing capacity of the autonomously intelligent machine. Thus, perhaps, the path of humanity might not get lost and confused in nebulous galaxies of data that, in the qualitative and quantitative volumes expected, risk exposing the surface of autonomous human thought to utter confusion. If today we are still strong in the certainty of an autonomy of thought, it is good to use this time to lay the (human) foundations of *Hyper-thinking*, a way that is totally new and predictably integrated with AI capabilities for the production of human thought.

To frame the scope of the debate, it is perhaps necessary to broaden our horizon of observation beyond the military field, considering digital transformation as a portion of the more articulated and broader framework of organisational innovation, whose starting point is always the individual, his culture and his

[137]Allen, J. R. (2018),*"Intelligenza Artificiale, una nuova era per la guerra"*, ISPI.
[138] Ibid.

awareness. The increasing centrality of the role of machines challenges the current anthropocentric vision, a vision in which humans - endowed with empathy and capable of handling exceptions and ambiguities - have, until now, governed every decisive choice, even and especially in the most complex scenarios. Can we imagine a future where machines will reach this level of discernment? A question that becomes even more crucial if we return to Defence territory.

As stated by the theologian Paolo Benanti[139], national security is a *"cross-cutting issue, in its physical, digital, social security meanings"* [140]. And technological innovation plays an increasingly central role in continuing to guarantee the highest level of security for citizens and territories. However, technology remains the tool, not the end. And it must be oriented and developed, *"having clear strategic goals and keeping the value course firmly in mind"*[141]. A course that calls into question the issue of the choice of rules, which must be commensurate with national perspectives and aspirations and, above all, with ethics itself, which can change at different historical moments and in different geographies of the world. These are global challenges that nations cannot tackle alone. In order to harness the potential and mitigate the risks associated with the application of AI, mutual influence should be harnessed with a view to international interconnection and interoperability. Today's confrontation is and will be with different actors and rules, in a legal framework still to be constructed, especially with regard to hybrid threats.

We are therefore faced with many variables, which combine to create a constantly changing context. Technological innovation is in fact, by its very nature, a dynamic element. So are the geopolitical scenarios and the different cultural evolutions that determine the choices of individual countries. The definition of rules and ethical principles is an articulated path, which is necessarily destined to go through such complexity in order to find common ground, a synthesis capable not of chasing, but of accompanying and fostering technological progress in every field. Hence the need for "permanent consultation" between all the actors that contribute to the definition of the rules of civilised living and those that govern, as a consequence, the rules of conflict: jurists, military, politicians, diplomacy and, last but not least, technologists - those who research, study, design and implement defence systems. Together, these different spheres must work for the same cause: to build a dynamic, circular and

---

[139] A Third Order Regular Franciscan, theologian, he works on ethics, bioethics and technology ethics. In particular, his studies focus on innovation management. The internet and the impact of the Digital Age, biotechnology for human enhancement and biosafety, neuroscience and neurotechnology

[140] Report of the 1st Defence Innovation Workshop 2021, *"Artificial Intelligence, Defence and Sustainability. Ethical, Legal and Psychological Implications",* April 2021 - Defence Innovation Centre (CID), with the collaboration of the European Institute for Innovation and Sustainability (EIIS)

[141] Ibid.

iterative logic that adapts the language of machines to human needs, taking ethical issues into account. A circular process that must never stop, that must consider ever more complex scenarios and that allows man to maintain a sense of limit in his idea of power and the use of force.

In conclusion, new technologies will continue to advance and reach new heights, new potentials, driven by future needs and continuous adaptation. Stopping this process is neither desirable nor imaginable. On the contrary, it is necessary to encourage and support generalised and equal access to innovations, setting in motion a virtuous and constant path that allows social and geopolitical bodies to *"metabolise"* technological progress within them. The economist Giacomo Becattini and the sociologist Luigi Burroni[142], in their analysis of Italian industrial districts, referred to this *"metabolisation"* as the capacity of territories - Institutions, Enterprises, Academia, Citizens - to know how to welcome, understand, manage and integrate technological innovation through *"decoding spaces"*, which are necessary to acquire "*the awareness of a complex and articulated process such as that of technological change"*. If we relate this reflection to the context we are analysing, the path to reach this awareness can only pass through a platform of permanent dialogue, involving all the key players that contribute to the construction of such scenarios - governments, armed forces, the academic and research world, companies - in a framework of multilateral confrontation. It is indeed on this shared awareness that we can base the moral axis and the perimeter of ethical principles within which to operate and delimit the use of new technologies in future Defence and Security scenarios.

[142] "*The industrial district as a tool for recomposing social knowledge*", Periodical "Sociology of Work" - Year: 2003, Issue 92

Conclusions:

- The implications of the development and application of disruptive technologies in defence are manifold and are changing the “rules of the game”. Speed, efficiency and accuracy of military operations, as well as the need for human control over AI-powered military capabilities, are at the heart of the issue in the current debate on artificial intelligence applications and national security.
- The traditional paradigm of international stability is challenged and will have to be renewed, taking into account weapons enabled by disruptive technologies and their effects on deterrence.
- The context of the near future requires the maintenance of a credible technological primacy, capable of fuelling an effective deterrence that induces hypothetical aggressors to carry out - before their respective initiatives - a cost-benefit assessment.
- The confrontation will take place with different actors and rules, in a legal framework still to be constructed, especially with regard to hybrid threats. Hence the need for “permanent consultation” between all the actors who contribute to defining the rules of civilised living and the rules of conflict, such as jurists, the military, politicians, diplomacy and those who research, study, design and implement defence systems.

Chapter V

LEGISLATION ON ARTIFICIAL INTELLIGENCE OUTSIDE THE EUROPEAN UNION: A COMPARATIVE ANALISYS

by Antonio Malaschini

SUMMARY: There's no general regulation on IA outside the European Union. China, United States, United Kingdom and Russia refer to current legislation in different fields; to the development, by way of interpretation, of existing rules for the protection of personal and social rights; to added production and trading standards; to guidelines and incentives. This is also probably caused by the dual nature, civil and military (including social control) of IA, that induces countries with strong international attitudes to restrain in disciplining and restricting this technology.

1. THE GLOBAL REGULATORY FRAMEWORK

There is currently no standardisation in the field of artificial intelligence that has global character and uniformity[143]. On the contrary, we can say that in the countries to which we shall refer, we will not find, even within them, a comprehensive regulation of the various aspects of AI: from the ethical to the research aspect, from the legal to that of its use, from that of potential risks to that of safeguards, from that of the protection of rights and freedoms to that of the relationship between public and private subjects, and so on. There is not even an agreed legal definition of what is meant by AI, and it is only in the most recent period that the opinion on the need for standardisation has been consolidated. Until now it was believed, and some still believe, that reference could be made to existing rules, such as those on cybersecurity and on the marketing and use of IT products; to the extension, also by way of interpretation, of the rules protecting personal or social rights; to production and use standards already designed for other goods; to the establishment of countless committees, commissions, and groups of experts producing valuable reports, guidelines, proposals, and resolutions almost never aimed at a complete discipline. Only the European Union, as we shall see later, made a reasonably organic proposal in April 2021. In the conclusions we will briefly mention what we consider to be the reasons for this regulatory delay. But let's see what the picture is at the moment in countries outside the EU.

[143] M.C. Buiten, "Towards intelligent regulation of Artificial Intelligence", in www.Cambridge.org/core, 15 April 2021.

2. CHINA

There is no legislation in China today that directly regulates AI, despite the fact that the country ranks among the leaders in the funding, research and use of this tool [144].

Although there is no specific source, there is nevertheless a set of provisions and guidelines that indicate with sufficient clarity the regulatory framework on AI: from the Cybersecurity Law (Csl) that came into force on 1 June 2017, to the National Industrial Information Security Development Centre's White Paper of January 2021, which outlined the lines of development of AI for the coming decades, to the recent Personal Information Protection Law of 20 April 2021 and the Data Security Law of 29 April 2021[145]. And lastly, the circular of the Council of State (the government) of 12 January 2022 on the development of the 14th Five-Year Plan and the role that the digital economy, AI and quantum information [146]should play in it.

The main point of reference in this regard, however, remains the "New Generation AI Development Plan" (Aidp) adopted by the State Council on 20 July 2017, which sets out the strategic lines to make China the most advanced country in the field of AI by 2030[147]. It is a document that has no immediate prescriptive value but, in the Chinese model, strongly conditions the behaviour of public and private actors in the sector.

The Plan is guided by a Strategic Advisory Board established in November 2017 and coordinated by the Ministry of Science and Technology. The government's strategy places the private sector and local governments at the forefront, and uses policy rather than regulatory instruments for its implementation. In the private sector, government-backed National Champions have been identified for specific projects in the field of Ai, e.g. Baidu for autonomous driving, Alibaba for smart cities, Tencent in the medical field. In return for their commitment, the companies will enjoy preferential public contracts, easier access to finance and market share reserves[148]. A similar system of incentives and facilitation allows local governments to implement decentralised initiatives.

Three areas are indicated in the Aidp as priorities. The first is that of international competition, and is linked to the military development of this dual

---

[144] Y. Luo, Z. Yu "An interview with Covington & Burling discussing Artificial Intelligence in China", in Lexology, 27 November 2020.

[145] G. Zou, H. Zhang, "China: a closer look at governmental and regulatory support of AI", in managingip.com, sponsored by Liu Shen IP, 3 March 2021.

[146] Xinhua Agency, www.gov.cn, 12 January 2022.

[147] In the translation by G. Webster, R. Creemers, P. Triolo, E. Kania at www.newamerica.org/cybersecurity-initiative/digichina/blog

[148] H. Roberts, J. Cowls, J. Morley, M. Taddeo, V. Wang, L.Floridi, "The Chinese approach to AI: an analysis of policy, ethics and regulation" in AI & Society, 36 (2021) pp. 59-77.

technology: military and civil. Rather than competing with the United States in traditional weaponry, China prefers to develop a policy that ensures it a decisive superiority in AI[149].

Also relevant is the programme for the use of AI for economic development: according to research by PriceWaterhouseCooper, China could secure a 26 per cent increase in GDP by 2030 through its use[150].

The third area of intervention, social governance, is the one that arouses the most interest in the West. Here, the Plan refers to the use of AI in the administration of justice, suggesting its use on the basis of the principle of "similar judgments in similar cases"[151]. It links its use to the so-called Social Credit System, which analyses individual and collective behaviour by awarding or deducting points on the basis of yardsticks established by the authorities. It suggests its adoption to understand group or individual dynamics and psychology (profiling) to improve school and work learning and evaluation systems, and for security and the management of anti-crime and anti-terrorist surveillance techniques, also through the discussed facial recognition[152]. The latter has seen in recent years, together with biometric recognition, a significant development underlined by a recent decision of the Supreme People's Court directing its use[153].

The Plan also refers to the need to define an ethical framework that could lead to a code of conduct to govern relations between AI and humans. On this point, it should be mentioned that in September 2021, the Ministry of Technology issued "Ethical Standards" through the National Guidance Commission for New Generation AI to integrate these principles into the AI lifecycle.

It is, however, as mentioned, a non-binding regulatory framework, which presupposes voluntary compliance on the part of public or private actors. But it is a voluntary adherence "with Chinese characteristics", as non-compliance may lead to sanctions of different kinds, such as those provided for, for instance, in the event of violation of the "spirit of the Personal Information Security Specifications", informed by the protection of that undefined "significant public interest", which is often the basis of Chinese control policies in the most diverse fields[154].

---

[149] D. J. Blackout, "Technology determines tactics: the relationship between technology and doctrine in Chinese military thinking", in Journal of Stratford Studies, 34, 3, 2011, pp. 355-382

[150] PriceWaterhouseCooper, Sizing the Prize, 2017

[151] S. Yuan, "AI assisted sentencing speeds up cases in judicial system", in China Daily, 18 April 2019.

[152] Y. Zeng, E. Lu, Y. Sun, R. Tian "Responsible Facial Recognition and Beyond", Institute of Automation, Chinese Academy of Sciences, 2020

[153] "Provisions of the People's Court on Several Issues Concerning the Application of Law in the Trial of Civil Cases Involving the Processing of Personal Informations Using Facial Recognition Technology", July 2021

[154] H. Yang, "China: the privacy, data protection and Cybersecurity law review", Edition 6, The Law Reviews, 2019

This was reaffirmed in March 2021 at the “two sessions”, the parallel meetings of the National People's Congress and the Chinese People's Political Consultative Conference, which, in launching the Digital China initiative, strongly emphasised the need for all actors in this field to comply with the often-referred to, but never well-specified, vital public interest.

Concluding this brief reconstruction, we can say that there is currently no legislation in China that specifically regulates this issue. References can be found, as mentioned, in the Cybersecurity Law, the Personal Information Protection Law[155], the Data Security Law and the Aidp. And in instruments such as the definition of production and utilisation standards, which aim to “steer” public and private initiatives according to government requirements[156]. It is an approach therefore not regulatory, but one of strong administrative discretion and sanctioning.

3. UNITED STATES

Even in the United States, there is currently no organic legislation on AI. The two opposing lines that characterise the debate on its regulation in all countries are clearly confronted here. On the one hand, the prevalence given in proposed regulations, as in Europe, to the protection of the rights of the individual, of the various freedoms, and of democracy itself in the face of an extremely invasive technology; and on the other hand, the military and strategic dimension of a “dual” instrument, military and civil, that cannot allow a global power such as the US to see its adversaries prevail. Hence the central role that the executive branch plays in regulation, which is more aware than others of the need to defend the country's international role.

[155] This law, as mentioned above, was passed on 20 April 2021 and came into force on 1 November of the same year. Among other things, it aims to regulate facial recognition and biometric recognition, on the other hand recognising their possibility of being used.

[156] In this regard, it is worth mentioning the adoption on 8 August 2021 by various public structures, starting with the Ministry of Industry and Technology, of guidelines for setting national and industry standards and organisational guidelines in the field of AI. See Y. Luo and Z. Yu: “An interview with Covington & Burling discussing artificial intelligence in China”, in Lexology, 5 January 2022.

3.1 REGULATORY INITIATIVES

In 2016, under the Obama presidency, two documents on AI came out. One is dated 12 October [157] and it concludes with 23 recommendations emphasising the role the government must play in fostering debate and coordination between the various public and private actors. The second[158], dated the following day, within the framework indicated in the previous document proposes a plan with six strategic objectives: stimulating long-term investment; developing effective models of human-IA collaboration; addressing ethical, legal and social implications; ensuring security and reliability; developing a shared data environment; and measuring and evaluating the technologies in question through common benchmarks. In the American tradition, the government aims to create a "fair playing field".

A direction more directed towards the strategic dimension of the use of AI was imparted by the Trump presidency. On 27 June 2018, the Undersecretary of Defence presented a memorandum for the establishment of a Joint Artificial Intelligence Centre, among all the Armed Forces, to engage with other governmental actors, industry and university and research institutions in order to "adapt AI technologies to Defence missions". Shortly afterwards, the 2019 budget identified these missions: to draw up a strategic plan; to accelerate the use of AI internally; to ensure governance to maintain the country's lead in these technologies, especially through training and recruiting talent[159]. Of note is the establishment in the same law of a National Security Commission on AI, with the task of making proposals to increase the country's competitiveness in the field of AI, with a particular focus on the need to protect America's technological advantage and military preponderance in confrontation with other states. We will refer to this commission later[160].

The following year, it was again the Presidency that relaunched the initiative with a memorandum and an Executive Order of 11 February 2019. Two indications among those that emerge from these documents seem to us to be particularly significant. The first is the need to ensure through scholarships and other incentives the development of numerically adequate and competent talent, while the second is the definition of a concrete plan of action to protect the country against "foreign competitors and opposing nations"[161]. Following the

---

[157] Subcommittee on Machine Learning and Artificial Intelligence del National Science and Technology Office (NSTC): "Preparing for the future of Artificial Intelligence".

[158] "The National AI Research and Development Plan" of the AI Task Force, a working group also within the NSTC

[159] John MacCaine National Defense Authorization Act for Fiscal Year 2019, sec. 238.

[160] Ibid, sect. 1051.

[161] Executive Order 13859 of 11 February 2019, Federal Register, vol. 84, no. 31, and sect. 7 and 8(a).

Executive Order, in August 2019, the National Institute of Standards and Technology then adopted a plan identifying nine areas of technical and non-technical standards to be followed, including ethical and social standards, governance and privacy protection among the non-technical ones[162].

Let us leave aside for the sake of brevity the further initiatives of the Presidency and the House of Representatives on our topic (including an update of the 2016 Plan mentioned above), and come to the most recent developments. In March 2020, the National Artificial Intelligence Act of 2020 was presented by a number of congressmen, which will then enter into force on 1 January 2021 as an amendment to the defence budget[163]. It is a bipartisan, executive and legislative proposal, and its aims are the latest direction in the field: to ensure American leadership in AI research and development; to make the United States the leading country in the reliable use of technology; to develop a capable workforce at all levels; and to coordinate research, development and use of these systems among civilian agencies, the Department of Defence and the Intelligence Community. To this end, a National Artificial Intelligence Initiative Office was established by the White House Office of Science and Technology, as a "focal point of government initiatives for Federal Departments and Agencies, industry, academia, non-profit organisations, professional societies, state governments, and any other parties the Initiative Office deems appropriate". An important role will be played by an Advisory Committee, appointed by the government, whose members will include "academic institutions, companies from various sectors, non-profit and civil society associations, including those protecting civil rights and the rights of the disabled, as well as federal research facilities expressing geographical diversity"[164]. On 8 September 2021, the Secretary of Commerce announced the formation of the Advisory Committee now mentioned.

To better clarify the picture that is emerging in the United States, it is perhaps also useful to mention a Federal Trade Commission (FTC) document dated 19 April 2021, i.e. when the European proposals that would be made official two days later were already known[165]. The FTC's invitation to developers and users of AI is to adhere to the standards already present in North American law to avoid prejudice based on race, colour, religion, nationality, gender, age, or receipt of public assistance. And so reference is made to the FTC Act of 1914, as amended in 2006 by the Safe Web Amendment; the Fair Credit Reporting Act of 1970 and

---

[162] y. Chae, "U.S. Regulation Guide: Legislative Overview and Practical Considerations", in The Journal of Robotics, Artificial Intelligence and Law, vol. 3, no. 1, January-February 2020. Let us recall, regarding the focus on "human-centred" values, that the US adopted the OECD's "Principles on AI" in May 2019, which place such values at the heart of the proposal.

[163] National Defense Authorization Act for Fiscal Year 2021, Sect. 5102, known as "Division E of the NDAA".

[164] Ibid, sect. 5104 (b).

[165] E. Jillson, "Aiming for truth, fairness and equity in your company's use of AI", Federal Trade Commission, 19 April 2021, www.ftc.gov/news-events/blogs.

its subsequent amendments; and the Equal Credit Opportunity Act of 1970 as amended. The intention is to suggest an ethical by design approach to manufacturers in order to comply with the standards contained in the standards mentioned above, which are currently considered sufficient to eliminate the biases referred to. It is a further address, for the time being, of opposition to a general discipline.

The reactions of the industrial and research worlds to this regulatory approach have been positive for the time being. The new guidelines are in fact seen as a point of reference that offers guidance without imposing onerous rules that could hinder the development of AI[166]. And here an attitude of criticism towards the new overall, stricter legislation being developed by the European Union is evident. It is a positive orientation favoured by the fact that on 8 June 2021, the Innovation and Competition Act allocated USD 80 billion for research in AI, robotics and biotechnology[167].

3.2 THE CONCLUSIONS OF THE NATIONAL SECURITY COMMISSION ON AI

As mentioned, the Defence Budget Act for 2019 established a National Security Commission on AI (NSCAI) to provide the Presidency and Congress with recommendations to "foster the development of artificial intelligence, machine learning, and associated technologies to comprehensively address the national security and defence needs of the United States"[168]. We find it useful to briefly describe the Commission's conclusions, which are perhaps the most illuminating document on the US strategy for AI.

Interestingly, the Commission includes, among others, a former CEO of Google, those of Oracle, In-Q-Tech and Amazon Web Services, a retired Marine colonel and former Undersecretary of Defence, the head of Google's AI division, and deans and professors from leading scientific faculties. A collection of experts, a total of 15, representing the vanguard of the scientific and productive world in the field of AI.

The conclusions are divided into two sections: defending America in the age of AI, and overcoming the technological challenge. The first discusses emerging threats, current and future defence issues, the role of AI in warfare scenarios, the

[166] S. Overly, M. Heikkilä, "China wants to dominate AI. The US and Europe need each other to tame it", in Politico, 2 March 2021, p. 6

[167] H. M. Lyon, F. A. Waldman, "Artificial Intelligence and Automated Systems Legal Update", Gibson Dunn and Crutcher LLP, in Lexology, 11 August 2021.

[168] National Security Commission on Artificial Intelligence, Final Report, March 2021.

risks of autonomous weapon systems, the relationship between intelligence and AI, and the human-machine relationship. The second section is devoted entirely to the technological challenge, primarily with China.

They are not optimistic conclusions and instead recognise that "America is not prepared to defend itself or compete in the age of AI. This is the harsh reality we have to face"[169]. And to "defend America" one must first recognise that Defence is lagging behind in the acquisition of new technologies; in the precedence still given to "heavy" structures such as ships, aircraft, mechanised vehicles and so on; in the lack of an effective strategy to recruit top talent; in obsolete analysis, acquisition, utilisation and control procedures; in the need to overcome an "industrial age" mentality. A delay that also manifests itself in the field of intelligence.

As far as the technological challenge is concerned, the second section sees the risk of losing out to China in the coming years. This is due to a lack of coordination within the government of the many initiatives on the subject; a lack of effective synergies with the private sector, academia and international partners; and the unresolved issue of "talent" to be trained and recruited. And again what is raised is the issue of intellectual property; that of standards and technologies associated with AI; and, last but not least, the American dependence on foreign countries (for more than 90 per cent) in the acquisition of microchips.

The overall sentiment is therefore one of great concern: "we must adopt AI to change the way we defend America". Hence a non-trivial organisational consequence: "competition over AI claims leadership of the White House".

Ethical aspects and the protection of freedoms and rights are not ignored, but "competitors are actively developing AI concepts and technology for military uses".

These are conclusions that are strongly influencing the American discussion and subsequent decisions on AI, and which seem to be dominated at the moment by concerns arising from global confrontation, primarily with China. As we will see shortly, these concerns are different from those that prompted the European Union to focus instead on how to protect its citizens in the use of a technology that strongly conditions rights, freedoms and democracy itself.

[169] Ibid p.1

4. UNITED KINGDOM

Brexit cannot fail to have an impact on the UK's AI policy. In fact, unless future agreements or transpositions are made, the rules being developed by the European Union cannot be applied. It is therefore necessary to examine today's framework, pending such possible understandings.

One of the characteristics of the UK is, in the field of AI, the relevant role of established scientific institutions such as the Alan Turing Institute and the Ada Lovelace Institute, which not only offer advisory support but are also the government's main interlocutors in the field[170].

It was precisely Turing that was referred to in January 2021 in the roadmap of the UK AI Council, a committee of experts from different fields, which presented the government with a document concluding with 16 recommendations for a national strategy[171]: develop public investment; strengthen the role of the Alan Turing as a national reference centre[172]; foster interdisciplinary AI education; increase public trust in tools that should be reliable, transparent, accessible and respectful of human rights; build an appropriate regulatory and sanctioning system; foster the creation of start-ups; promote international cooperation; develop health-related applications; but also increase military and homeland security through AI. A programme, as you can see, not devoid of ambition.

To this document we must add the recommendations in April 2021 of the Centre for Data, Ethics and Innovation (Cdei), which reiterates the need for a reliable model capable of reassuring citizens about this new technology[173]. Recommendations updated in December of the same year to emphasise the role that AI could play in three areas: recruitment and management of workers, in order to eliminate discriminatory practices; transport and logistics, also to increase energy efficiency; and the education sector[174].

Again, on 22 September 2021, a 10-year plan (National AI Strategy) with three pillars was presented in Parliament by the Minister for Digitalisation: long-term investment, ensuring benefits for all sectors and areas of the country, and ensuring effective governance[175]. A number of medium-term objectives are set

---

[170] See most recently the Ada Lovelace Institute's report “Regulate to innovate” on the regulatory tools to be used, which should start with a clear definition of what is meant by AI systems: see Tom Whittaker and Liam Edwards, “Improving AI governance in the UK”, Burges Salmon LLP, in Lexology 8 December 2021.

[171] www.gov.uk/government/publications/ai-roadmap

[172] It is worth mentioning here that in January 2022, the Turing was chosen by the government as the “pilot” for a project for a new standard-setting centre for AI. See Natalie Donovan: “New AI standards hub launched in UK”, Slaughter and May, in Lexology, 14 January 2022.

[173] See Natalie Donovan, "New AI assurance roadmap published - is this the route to safe AI?" Slaughter and May, in Lexology, 9 December 2021.

[174] Sam Morrow and JP Buckley, “Artificial Intelligence Update”, DWF LLP, 17 January 2022.

[175] https.publishing.service.gov.un/government/uploads/systems/uploads/attachment_data/file/1020402/National_AI_Strategy_-_PDF_version.pdf

out in the plan: to include AI among the subjects of training courses (bootcamps) organised by the Ministry of Education for sectors with specific technical skills at the request of entrepreneurs; to prepare guidelines for the responsible collection and use of data; to open a public consultation (which was then actually held); and to maintain a central role for intellectual property in order to promote creativity and innovation[176].

A different approach from that of the European Union is currently emerging, even on issues of common interest such as the very definition of AI, data protection, transparency, bias, and criteria defining responsibilities. A cross-sectoral and not an all-encompassing EU regulation, with sector-specific codes of conduct and policy-driven guidelines that will take into account the peculiarities of the different fields of action[177]. In line with the recommendations of the House of Lords, which already in 2018 considered that "specific legislation on AI would be inappropriate at this stage"[178].

And on the subject of the House of Lords, it is worth noting the significant correspondence of one of its documents dated 18 December 2020 with the worrying conclusions of the US National Security Commission on AI that we have quoted above. In its report on the government's activities in this field, significantly entitled "*No room for complacency",* it points out the delays, and the strategic implications of those delays, that strongly influence the framework of international relations. Once again, in the dual world of Artificial Intelligence, global confrontation with antagonistic countries proves to be central[179].

---

[176] This is the issue dealt with in the well-known "Dabus" case, the appeal filed in 2018 by S. Thaler against the decision of the British Intellectual Property Office (Ukipo) to reject his application for two patents relating to allegedly autonomous inventions resulting from an AI system. In an alternation of judgements, the ultimate decision is currently against recognition. See Lisa Page, Thomas Kirby and Shaan Mehra, "UK Court dismisses Dabus - an AI machine cannot be an inventor", Penningtons Manches Cooper LLP, in Lexology 14 December 2021; Mike Pierides and Katherine B. O'Keefe, "AI and UK Intellectual Property Consultation: Copyright and Patents", Morgan, Lewis & Bockius LLP, in Lexology 14 January 2022; Xin Hu Rasmussen and B. Vinti, "Update on artificial intelligence as a patent inventor", Proskauer Rose LLP, in Lexology, 18 January 2022.

[177] J. McDonalds - Charles Russell Speechlys: "Regulating AI - the impact of two recent proposals: the UK's National AI Strayegy and the EU's proposed Artificial Intelligence Regulation", in Lexology, 27 October 2021; and Tom Whittaker and Liam Edwards, "Improving AI governance in the UK", cit.

[178] Select Committee on Artificial Intelligence, *AI in the United Kingdom: ready, willing and able?,* House of Lords paper, 100.

[179] Liason Committee on Artificial Intelligence, *AI in the UK: no room for complacency,* 18 December 2020, session 2019-2020, HL papers 196.

5. RUSSIA

In the opinion of analysts, Russia is regarded as an “outsider” in the field of AI[180]: a term that, in a less than benevolent interpretation, could be translated as “the outcast”. It means lagging behind its strategic competitors, primarily the US and China. As early as 2017, however, Vladimir Putin declared that “whoever becomes the leader in artificial intelligence will rule the world”. And in October 2019 he regulated the “National Strategy of the Russian Federation for AI” with his decree[181]. The document starts with the recognition of the usual principles: protection of human rights, fundamental freedoms, security, transparency and so on. The strategy seems to be based on the prevalence of AI research and development in the economic and financial field, in services and in the health sector, and does not expressly mention security and defence applications: a singular omission given the development and application of AI in Russia in the two sectors now mentioned.

An important point in this strategy is the creation of a national database into which data can be stored, with priority access to the data for public authorities. The predominant role played by the state in this “Strategy” is underlined by the fact that in recent years 67% of funding for AI research and development in Russia comes from the state budget, while private resources account for 79% in China and 77% in the US[182].

The implementation of the strategy, which is projected to 2030, is delegated to a number of government organisations with a coordinating role played by the Digital Development Commission under the Ministry of Digital Development, which is responsible, among other things, for coordinating the 13 projects launched by the government in 2018 for the modernisation of the Russian economy.

Interestingly, the National Strategy was prepared with the significant contribution of Sberbank, Russia's largest bank, 52% owned by the state, which intends to become a “company for artificial intelligence”, reserving an immense amount of data for itself and the government[183].

The fact that the National Strategy does not explicitly refer to military uses does not, of course, mean that these are excluded from the programme: this is shown, among other things, by the fact that Putin's 2019 decree, in increasing the appropriations for AI research from USD 1.3 billion to USD 6.1 billion, does not

---

[180] J. Nocetti, “The outsider, Russia in the race for AI”, in Russie. Nei. Reports, no. 34, Ifri, December 2020.

[181] Decree of the President of the Russian Federation on the Development of AI, 10 October 2019, in http://publication.pravo.gov.ru

[182] “Leading indicators for science and technology”, OECD, 2019.

[183] E. Tofaniuk, N. Uskov, “German Greg: Transforming Sberbank is a continuing process”, in Forbes, Ru, 22 November 2019

distinguish between civil and military uses, and it seems credible that these funds will also increase those already provided for in the defence budget[184].

Once again, therefore, we are faced with a model with a decisive predominance of the state in terms of standardisation, direction and control. It is a trend that, albeit with significant differences, now seems to be consolidated in China, the United States and Russia, which is contrasted, as we shall now see, by that of the European Union.

6. A FIRST COMPARISON

One element unites, albeit with marked differences, the attitudes of the countries we have examined so far. And that is what they are not currently willing to promote internally as a comprehensive discipline on AI. Instead, with an approach that could be called "regulation by design", they are setting standards for manufacturers rather than rigid rules whose compliance will be verified *ex post*[185].

Alongside this, the central role that executives have carved out for themselves to guide the process should be emphasised. Also of extreme importance is the focus in these countries on aspects of global politics and strategic confrontation, which place the applications of AI in the military and security field in the foreground: to the detriment of the concerns that this medium raises for the protection of privacy, for the defence of personal, economic and social freedoms, and for the process of democratic participation and decision-making.

Certainly on this last aspect, there are extremely significant differences. For example, the obsessive use of tools such as facial recognition and predictive techniques for the purposes of control and social repression, as happens in China, certainly cannot be compared to the American system, which retains political, legal, press freedom and economic instruments that can intervene in defence of the oft-mentioned freedoms. Just as the unscrupulous use of AI with international political objectives capable of influencing the electoral competition, but also the economic and social stability, of opposing countries as seems to be the case in Russia is quite different from the security concerns of the UK[186].

---

[184] On this topic, P. Luzin, "Artificial Intelligence in the Russian Army", in Riddle, 19 March 2021.

"[185] B. Docquir, R. Garcia del Poyo, X. Pican, T. Quinn: "Legislators worldwide move to adopt regulation by design", Osborne Clarke, in Lexology, 24 January 2022

[186] On the topic of the use of AI in cyber-warfare between nations, particularly but not only by Russia, see: "This is how they tell me the world ends", by Nicole Perlroth, Waterstones, 2021, which focuses on the use of so-called "zero days" activated through AI in strategic confrontation and financial, industrial and commercial struggles.

There is no doubt, however, that in the countries we have examined so far, all with long traditions of strategic prevalence in global confrontation, the concerns that we may for simplicity's sake describe as “ethical” appear, albeit with varying degrees of intensity, to be recessive compared to those of military defence.

It is a dilemma that history has put forward more than once. Think, for example, of the fierce confrontation in the 1950s and 1960s over the use of nuclear energy for military and strategic purposes or for peaceful purposes, a confrontation that has never been definitively resolved and which is even being re-proposed these days with reference to industrial use only.

We will return to this aspect in the conclusions, after examining the legislation being drawn up in the European Union, which seems to differ from that of the countries examined so far in that it is more sensitive, we repeat again, to the issue of the protection of rights and freedoms in the face of a phenomenon potentially capable of undermining them.

Conclusions:

- There’s no general regulation at the moment on IA.
- China, United States, United Kingdom and Russia focus more on guidelines, incentives and standards, than on punctual legislation.
- The military and social control aspect of AI strongly conditions the politics of these countries

Chapter VI

LEGISLATION IN THE EUROPEAN REFERENCE FRAMEWORK AND THE ITALIAN CASE

by Antonio Malaschini

SUMMARY: Unlike other countries, the European Union is following the path of a proposal for comprehensive regulation of AI. The guiding principle is that of a "human-centred" AI, which puts the individual and his or her rights, both personal and social, at the centre of regulation. Thus, it tries to eliminate the "risks" that this instrument can entail for individuals and society. It is an orientation followed, albeit with some delay, also by Italy.

1. INTRODUCTION

After the positive experience of the GDPR in the field of data protection[187], it was natural for the European institutions to focus their attention on the need to guide the relatively new, and controversial, phenomenon of AI also, as we will see shortly, from a perspective that is not only internal to the European Union.

Without wishing to retrace in detail the process that would later lead to the adoption by the European Commission of its proposal for a regulation, we cannot fail to point out some of the stages along this path, starting in April 2018 when the Commission, accepting the invitation of the European Council, hypothesised a strategy that would place "citizens at the centre of the development of AI, a human-centred AI"[188]. This will remain the inspiration for all the subsequent development of the Union's bodies, differing, as we have repeatedly recalled, from that of the countries outside the Union: more focused on reassuring its citizens and protecting their rights and freedoms as a priority; less concerned with the consequences at a strategic level of a policy of control in the field of research and development.

A "Co-ordinated Plan" for AI was then presented as early as December 2018, indicating the framework within which the member states were to include their national initiatives[189]. The inspiring idea was to foster a reliable technology that would reassure and prepare society for the impact of AI.

However, we cannot fail to mention some of the decisions subsequently taken by the Commission: the establishment in April 2018 of the High Level Expert Group (AI Hleg), with advisory and steering tasks; the AI Alliance in June of the same year, a discussion forum open to civil society; the AI Watch, in December 2018, to monitor the development, application and impact of AI in Europe. The first AI Hleg report, published in April 2019, pointed to four ethical principles that should guide AI systems: respect for human autonomy; harm prevention;

[187] This is the General Data Protection Regulation (Regulation No. 2016/679) of 27 April 2016 on the protection of individuals with regard to the processing of personal data.
[188] Com (2018) 237 final
[189] Com (2018) 795 final

fairness; explainability[190]. The report will subsequently be updated in July 2020, and will form the basis for the White Paper on AI that the Commission will publish on 19 February 2020: "A European Approach for Excellence and Trust"[191]. We will find many of the conclusions of the *White Paper* within the draft regulation. A further Commission document of 23 July 2020, the*Inception Impact Assessment,* takes stock of the proposed initiatives and their foreseeable effects[192]. According to this paper, the economic benefits of public intervention in the sector will only come from an increase in public confidence in the system.

Space does not permit us to recall the initiatives on the subject also taken by the other European institutions: Parliament, European Council, Council of the European Union. We would just like to mention what is called the Parliament's "October Framework on AI" in 2020, which sets out a series of ethical principles that we will find in the subsequent Commission proposal: security, transparency and accountability; guarantee against prejudice and discrimination; social and environmental responsibility; privacy and data security; identification of those high risks to be strictly regulated, on which human control cannot fail to be decisive[193]. In the "Framework", Parliament also addresses the issue of the military use of AI, reiterating the need for absolute human control of military systems that use it and calling for a ban on lethal autonomous weapons that make use of AI[194].

But let us finally come to the Commission's draft regulation[195].

2. THE EUROPEAN COMMISSION'S DRAFT REGULATION

On 21 April 2021, the European Commission will submit a draft AI regulation to the Parliament and the Council[196]. Let us briefly look at the essential points.

To begin with, those to whom the Regulation is addressed are providers who market or provide AI services in the Union, regardless of whether they are

---

[190] European Commission, Science for Policy Report, *AI Watch. Artificial Intelligence in Public Services,* 2020.

[191] Com (2020) 65 final. On this point see V. Litvinets, "A summary of the European Commission White Paper on AI", in medium.com, 9 June 2020.

[192] "Proposals for a legal act of the European Parliament and the Council laying down requirements on AI", 23 July 2020, Cnect. A. 2

[193] European Parliament Resolutions of 20 October 2020.

[194] G. Olivi, "The starting package", Dentons LLP, in Lexology, 23 November 2020.

[195] The draft regulation does not exhaust the framework of the Commission's interventions on the topic of AI. At the same time as the Regulation, an update of the 2018 Coordinated Plan on AI, mentioned above, was presented; and a new Machinery Directive to facilitate the integration of AI systems into machinery, and one onproduct liability, is in the pipeline.

[196] *Proposal for a Regulation of the European Parliament and of the Council: Laying Down Armonised Rules on AI* (*Artificial Intelligence Act*) *and Amending Certain Union Legislative Acts,* 21/4/2021, Com (2021) 206 final.

established inside or outside the Union; users within the Union; and suppliers or users of AI systems in a third country, if the system product is used in the Union. As with the GDPR, the rules therefore have a scope reaching beyond the EU boundaries.

The Regulation also proposes a definition of AI systems that is intended to be technologically neutral and such that it can be updated in the future by means of a special delegation to the Commission itself[197].

2.1 THE RISK CRITERION

The model proposed by the Commission is based on “risks”: the higher the risks, the stricter the rules. Starting with practices that generate such high risks that they are even prohibited.

Extreme risk practices are therefore prohibited when they use subliminal techniques to distort a subject's behaviour*(*dark patterns*,* such as the use of unperceivable sounds to influence behaviour); exploit a person's vulnerability in relation to age or physical or mental disability; are used by the authorities to assess or judge a subject on the basis of his or her personal or social characteristics (social scoring); and prohibit facial recognition practices in public places by the police, unless they are necessary in a limited timeframe to search for victims of a crime, to prevent attacks or to discover a subject accused of serious crimes.

Instead, there will be high risks when two conditions are met. The first is that AI systems are used as safety components of a product regulated by previous EU directives, such as machinery, medical instruments, toys, personal protective equipment; cars, lifts, aircraft[198]. The second is that they are self-sufficient systems, operating independently of other units (stand-alone), whose safety component is subject to a prior conformity assessment by a third party with a view to their entry onto the market.

Practices that are capable of influencing the exercise of the fundamental rights listed in Annex III of the proposal will always be considered high risk[199].

---

[197] Article 3 and Annex I define AI as software developed with one or more techniques or approaches (including machine learning and deep learning) that in relation to human-defined goals can generate different results that in terms of content, predictive capacity, recommendations or decisions can influence the environment with which they interact.

[198] See Regulation, Exploratory Memorandum, p. 13. and DLA Piper, “The World's First Artificial Intelligence Act”, in Lexology, 30 April 2021.

[199] These rights include those relating to identification and profiling of persons; management of critical infrastructure; education and training; recruitment and evaluation of personnel; access to essential public services; public order; migration, asylum and border control; administration of justice and the democratic process.

In order to be permitted, high-risk systems must first undergo compliance validation; then adopt appropriate risk management models; and provide for continuous estimation and evaluation processes. Up-to-date and verifiable technical documentation will always be required and they must be designed and developed in such a way as to ensure transparency, accountability and effective and continuous human control.

Validation is essentially based on self-assessment, with the exception of remote biometric identification systems and public infrastructure networks, for which validation by a third party is required. We will shortly see the penalties for non-compliance. In order to facilitate research even in high-risk areas, the creation of "regulated test environments" (sandboxes) is permitted[200] while complying with the planned safeguard regulations.

Systems that are not regarded as prohibited or high risk, on the basis of the above-mentioned characteristics, are not subject to different regulation. Most of the systems in use in Europe today, such as video games and spam filters, fall into this category. The Commission's hope is that voluntary codes of conduct will be adopted by companies, which, perhaps adopting the same caution as for high risks, will increase public confidence in these new technologies and, consequently, their diffusion in the market. An approach that is partly reminiscent of the "regulation by design" we mentioned earlier.

2.2 GOVERNANCE E SANCTIONS

A European AI Board has been established to assist the Commission in coordinating national policies, also with a view to uniform implementation.

The Board will be composed of the heads of the various national supervisory authorities and the European Data Protection Supervisor, and will be chaired by the Commission itself. At Member State level, national authorities are foreseen to ensure compliant enforcement, and a Supervisory Authority with market surveillance tasks.

To facilitate the Commission's work, a European database for "high risks" arising from stand-alone systems will be set up, to which providers will be required to register their systems.

National supervisory authorities will have the power to monitor and investigate compliance with the obligations and requirements. This *ex-post* control gives

[200] Davis Wright Tremain, in Lexology, 6 May 2021.

controlling entities the power to intervene, in particular if systems generate unexpected risks.

As far as sanctions are concerned, violations of the rules on prohibited practices or those on required data are punishable by an administrative fine of up to euro 30 million or a sum of up to 6 per cent of the offending company's annual total revenue. Violation of rules other than those now mentioned may instead lead to a penalty of up to euro 20 million or 4% of the company's revenue. The European Data Protection Supervisor is authorised, in the event of non-compliance, to impose sanctions on the institutions and agencies of the Union responsible.

As far as the next procedural steps of the proposal are concerned, once the planned public consultation is over, the usual process between Parliament and the Council of the European Union will begin, and it may be useful here to reach consensus through the "trialogue" with the Commission[201]. Once the legislative process has been completed, companies will be given time to adapt to the new requirements and the new rules will start to apply 24 months after the Regulation comes into force.

In the meantime, it will be up to the Member States to consider the changes required by the Regulation to their domestic legislation.

2.3 THE INITIAL CRITICAL OBSERVATIONS

The Commission's proposals, as was to be expected, stimulated a debate in which several critical voices emerged. Apart from those that refer to the imprecise, if not deliberately ambiguous language of the proposed rules, the criticisms put forward can be roughly divided into two categories: those that consider that the new rules restrict the exercise of various rights of citizens and the principles of freedom and equality; and those that, on the contrary, see them as an excessive obstacle to research and development of new technology, resulting in significant costs for the required adaptations, which are particularly burdensome for small and medium-sized enterprises[202].

This highlights the difficulty for companies, on the basis of the proposed rules, to balance the demand for transparency with the protection of intellectual freedom (IP) and trade secrets[203]; the obscurity of the concept of "explainability",

---

[201] The "trialogue" is an informal EU procedure involving Parliament, the Council and the Commission to facilitate an agreement within the ordinary legislative procedure.

[202] *The EU's approach to AI,* International Institute for Strategic Studies, vol. 27, Comment, 24 September 2021

[203] V. Karen, L. Neumann, M. Tierney, H. Bonaccorsi, M. White, "European Commission proposed Regulation on AI", Lexology, 21 June 2021.

which is supposed to make processes and algorithms comprehensible, but whose clarity sometimes escapes the developers themselves[204]; the vagueness of the criteria for distinguishing between high and low risks; the difficulty, for accountability purposes, of identifying a final responsible party; the centrality given to providers, now overtaken by a technological development that seems to favour open source models over traditional "suppliers"*;* the failure to take sufficient account of Ai-informed algorithms used in social media, search, online sales and applications for mobile devices and operating systems [205]as high risks.

Significant, then, are the criticisms expressed on 18 June 2021 in a Joint Opinion, albeit non-binding, of the European Data Protection Board (EDPB), which firstly calls for the prohibition of "any automatic recognition of human physiognomy in publicly accessible places, in any context". It then proposes a ban on AI systems that categorise subjects in relation to ethnicity, gender, political or sexual orientation. Again, pointing out the lack of autonomy vis-à-vis the Commission of the envisaged supervisory authorities, as well as the lack of coordination between the new regulation for AI and the existing regulation for data protection [206].

It is also worth mentioning, in a sense that does not always coincide with the comments above, that on 2 November 2021, a concerned and critical report by the Special Committee on AI in a Digital Age (AIDA) of the European Parliament was presented by rapporteur Axel Voss, which emphasised the Union's considerable backwardness in the research, use and regulation of AI, particularly in comparison to the United States and China. The report, among other things, calls for the exclusion from the limiting principles of legislation of the use of AI in connection with military and security activities in a world in which several countries are making firm use of this tool in the field of global confrontation and internal control[207]. By and large, he was less optimistic about the Commission's assessments and proposals.

Finally, on 29 November 2021, the Council of the European Union presented a draft compromise (Compromise Text) on some of the contested points. In particular, with regard to the definition of "Artificial Intelligence" itself, a text considered to be clearer is also proposed in order to prevent the inclusion in the new discipline of more traditional software systems that are not normally considered AI. In addition, changes are suggested that are deemed to better protect the rights of end-users by clarifying the concept of provider; national security is excluded from the scope of the proposed regulation, as under the

---

[204] G. Olivi, C. Bocchi, "Regulating AI in the European Union: top 10 issues for business to consider", Lexology, 1 July 2021.

[205] J. Buyers et al. "Debate continues over the pros e cons of regulating AI", Osborne Clarke, in Lexology 27 July 2021

[206] V. Horgan "Artificial Intelligence update", Bird & Bird LLP, in Lexology, 12 July 2021.

[207] European Parliament, Special Committee on AI in a Digital Age, Draft Report, 21/112021 (2020/22661 (INI). On the aspects of military use, see no. 154, p. 32.

Treaties it falls under the exclusive responsibility of the Member States; and the regulation for activities related to Research and Development is eased, in order to foster innovation processes, by excluding systems used exclusively for scientific research. Other important proposed amendments concern the extension of the ban on social scoring to private actors; the amendment to Annex III of the proposal insofar as it is proposed that AI systems that endanger the protection of the environment should also be considered high risks; and the provision of an Annex IV on accompanying documentation for high risks, in order to better ensure that human control measures are taken on AI systems[208].

What I have just said emphasises that, in the face of legislation that is not only relatively new but also poses ethical and social problems alongside the scientific, economic and production ones, the path of the global approach followed by the Union is undoubtedly fascinating but not without significant and complex problems, which will require a realistic and non-ideological approach in order to make the new discipline accepted by all: citizens, companies and states

3. ITALY

As mentioned earlier, in 2018, the European Commission's Coordinated Plan on AI required member states to define their own national strategies. The task was entrusted in Italy to the Ministry of Economic Development, whose owner announced an Italian AI strategy in July 2018 and then appointed a group of experts in January 2019.

A document was produced in which 82 proposals were made, which was then submitted for public consultation in September 2019 and updated in February 2020. In the document, which was never officially released, in addition to examining AI from a scientific, ethical, industrial and social perspective, the issue of governance was addressed: from the creation of a steering committee to the identification of a national institute for AI to assist and advise the government on these issues[209].

The events surrounding the emergence of the COVID-19 pandemic and subsequent political developments have probably contributed to putting the need to conclude this document on the back burner. With the Draghi government (13 February 2021), as far as we are concerned, the competences of the various ministries on the subject of AI are being modified, with the strengthening of the Ministry for Technological Innovation, which has absorbed, together with the

[208] William RM Long, Francesca Blythe and Subhalakashmi Kumar; “EU Council Publishes Changes to Artificial Intelligence Act Proposal”, Sidley Austin LLP, in Lexology, 18 January 2022; Adam Finlay and Catharine Walsh, “Council Publishes Proposed Amendments to Draft AI Regulation”, McCann FitzGerald LLP, in Lexology, 23 December 2021.

[209] "Proposals for an Italian Artificial Intelligence Strategy", in mise.gov.it/images/stories/documents

Ministry for Universities and Research, several powers that were previously the responsibility of the Ministry for Economic Development; thus favouring the establishment of an Inter-Ministerial Committee for Digital Transition between the institutions now mentioned.

On 7 October 2021, the working group set up in July of the same year by the ministers concerned delivered its recommendations to the government, which were subsequently approved on 24 November by the Council of Ministers[210]. The document outlines six objectives, eleven priorities and three focus areas, to be realised by 2024.

The three areas of focus are: enhancing skills and attracting talent; increasing funding for advanced AI research; and stimulating the adoption of AI and its applications. The objectives are to strengthen frontier research in AI; reduce the fragmentation of AI research; develop human-centred and reliable AI; increase AI-based innovation and technology development; develop AI in the public sector; and create, retain and attract AI researchers.

The priority sectors are industry and manufacturing; education system; agribusiness; culture and tourism; health and wellness; environment, infrastructure and networks; banking, finance and insurance; public administration; smart cities, areas and communities; national security; and information technology.

One point that should perhaps be explored further is that of governance. The document refers to a permanent working group within the Interministerial Committee “to direct, monitor and evaluate the implementation of this strategy and its subsequent interactions as well as to coordinate all AI policy actions in the future”. It is then stated that 'this implies the possibility of involving other institutional actors, universities and research centres as well as representatives of the private sector”.

This is certainly a positive solution insofar as it broadens the field of stakeholders. But the development of AI in our country requires, in a framework of increasingly complex global competition, a stronger and more authoritative point of reference. As was recently done in the area of cybersecurity, with the establishment, alongside an Inter-Ministerial Committee with advisory functions, of an operational structure, the National Cybersecurity Agency, with a Director appointed by the Prime Minister at the head, to whom he is the direct contact. And the identification of advanced research centres with not only an advisory but also a steering role should be encouraged, as in the UK with the Turing Institute and Ada Lovelace Institute.

As proof of the interest that seems at last to be accompanying reflection in Italy on AI, it is worth remembering how Prime Minister Draghi, in his communications to the Houses of Parliament on 20 October 2021, indicated as

[210] “Strategic Programme for Artificial Intelligence 2022-2024”, in https://assets.innovazione.gov.it

the government's objective on AI “to promote experimentation and, especially, to make its direction more secure and transparent. In the same way, we must increase public confidence in these new technological solutions”[211].

4. CONCLUSIONS

We have already indicated the reconstructive thread between the different regulatory directions. On the one hand, there are countries, such as the European Union, that place a “human-centric” vision at the centre of their actions in the field of Artificial Intelligence, hence the need to protect first and foremost the rights, freedoms of the individual, and democracy itself. Also accepting the risk of limits to research, production and use of AI systems.

On the other hand, countries that, also due to their past or current global dominance, are more attentive to the strategic dimensions of this instrument. Of course, they do not deny the need to protect the fundamental rights in question, but they reject at the moment a comprehensive approach such as the European one, preferring sectoral and guiding interventions rather than general standardisation that does not restrict research and use in both the civil and military fields.

There are, of course, as we have seen, profound differences between the US and UK approaches to this issue compared to those of China and Russia. In the latter countries, in addition to the need not to restrict civil and military research, there is a definite use of AI for internal security and social control purposes. And to a deemed use in the disruption of production, energy, financial, communication and electoral systems of adversary countries[212].

In this complex framework, it appears necessary in the first place, though not easy, to create an international discussion table where allegedly common ethical and social, research and production concerns can confront each other. And some initiatives already taken by the United Nations go in this direction, such as *The*

[211] Within this framework of work in progress, it is worth mentioning the very interesting document, albeit dating back to March 2018, from the Agenzia per l'Italia Digitale (Agid), “Artificial Intelligence at the Service of the Citizen”. The aim here was to analyse the impact of AI in Italian society, particularly in the public administration.

[212] See Note 42

UN's Secretary General's Strategy on New Technologies and the *Systemwide Strategic Approach and Roadmap for Supporting Capacity Development on Artificial Intelligence*[213].

We are not so naïve as to believe that the problem of the use of AI, which is already complex at the level of individual states, can be solved in this way. However, even the mere beginning of an international discussion on the civil aspects in this field would be a positive thing in itself. For the military, the subject of its regulation brings us back to the long, arduous and often unsuccessful attempts that have accompanied and still accompany the use of atomic energy. Attempts which have nevertheless prevented the atomic holocaust for the time being and which could also bear fruit in our field.

One last note before concluding. So far we have been talking about states, their internal decisions, their strategic confrontation. However, there are other relevant players.

All the countries mentioned above, and of course others, are indebted in the field of AI research and utilisation to companies whose size exceeds that of individual, major nations. A new acronym, MAAMA, unifies five American companies that now dominate the most advanced technology sector: Microsoft, Apple, Amazon, Meta and Alphabet. The same is happening in China, with Huawei, J.D. Com, Alibaba, Tencent and China Mobile. These companies, and of course others similar to them in other countries, dominate the field of research and use of AI-based tools. For example, if data are today the mines of knowledge and thus of development, we are in this field facing a situation of decided oligopoly[214]. This is not the place to examine this element in depth. We limit ourselves here to pointing out a problem that in China, for example, with Alibaba but not only, has led that government to adopt harsh measures of repression and direction.

These are not measures that we consider desirable or possible here per se. To ignore these players when talking about standardisation would, however, not only be a mistake but a risk, the consequences of which would perhaps render many of the initiatives that governments are now trying, with difficulty, to adopt vague and useless.

[213] United Nations, 2018, in www.un.org/en/new technologies/pdf/
[214] On this topic, see The Economist, 22 January 2022.

Conclusions:

- The European Union is the only country to propose a comprehensive regulation on IA.
- Its framework is based on the protection of the personal and social rights of the individuals, and on the elimination of the “risks” connected to the use of IA.
- Italy is following, albeit with some delay, the European approach

Authors

- Enrico Prati, Associate Professor in Condensed Matter Physics Theory - Università Statale di Milano
- Andrea Gilli, Senior Researcher - NATO Defense College
- Lucrezia Scaglioli, Research Assistant - NATO Defense College
- Enrico Savio, Chief Strategy and Market Intelligence Officer - Leonardo
- Enrico Comin, Strategic Planning Analyst - Leonardo
- Antonio Malaschini, Advisor to the Italian Minister of Economy and Finance