Presidency of the Council of Ministers

COMITATO NAZIONALE PER LA BIOETICA

COMITATO NAZIONALE PER LA BIOSICUREZZA LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA

DEVELOPMENTS IN ROBOTICS AND ROBOETHICS

17 July 2017

Members of the working group

For the Italian Committee for Bioethics:

Prof. Lorenzo d'Avack, Professor of Philosophy of Law, University 'Roma Tre', Deputy President of the Italian Committee for Bioethics (coordinator)
Prof. Riccardo Di Segni, Chief Rabbi of Rome, Vice President of the Italian Committee for Bioethics (coordinator)
Prof. Laura Palazzani, Professor of Philosophy of Law, Lumsa University of Rome, Vice President of the Italian Committee for Bioethics (coordinator)
Prof. Salvatore Amato, Professor of Philosophy of Law, University of Catania (coordinator)
Prof. Luisella Battaglia, Professor of Moral Philosophy and Bioethics, University of Genoa (coordinator)
Prof. Carlo Casonato, Professor of Comparative Public Law, University of Trento
Prof. Marianna Gensabella, Professor of Moral Philosophy, University of Messina
Dr. Maurizio Benato, Delegate of the National Federation of MDs' and Dentists' Colleges
Dr. Carlo Petrini, Delegate of the National Institute of Health

For the Italian Committee for Biosafety, Biotechnology and Life Sciences

Prof. Andrea Lenzi, Professor of Endocrinology, University of Rome "La Sapienza", Chair of the Italian Committee for Biosafety, Biotechnology and Life Sciences (coordinator)
Prof. Carlo Caltagirone, Professor of Neurology, University of Rome 'Tor Vergata' (coordinator, also a member of the Bioethics Committee)
Prof. Roberto Cingolani, Scientific Director of the Italian Institute of Technology (coordinator)
Prof. Roberta Siliquini, Director of the Superior Health Council (coordinator)
Prof. Antonio Amoroso, Professor of Medical Genetics, University of Turin
Prof. Pier Angelo Morandini, Lecturer in agricultural and environmental biotechnologies, University of Milan
Prof. Pier Franco Pignatti, Professor of Medical Genetics, University of Verona

For the support unit

Dr. Agnese Camilli (coordinator)
Ms Lorella Autizi
Dr. Patrizia Carnevale
Dr. Raffaella Falco
Dr. Andrea Giovannelli
Dr. Dorothea Karic
Dr. Stefania Mariotti
Mr. Angelo Rocchi
Dr. Carlo Santoro

CONTENTS

Introduction

This opinion on the scientific, technological, ethical and legal implications of robotics and roboethics has been prepared by a mixed group comprising members from the Italian Committee for Bioethics and from the Italian Committee for Biosafety, Biotechnology and Life Sciences.

It begins with a preliminary definition of “robot” and “artificial intelligence” and goes on to outline the possibilities, potential and limitations of the emerging new technologies in relation to robots with and without a mechanical body and with and without intelligence. At the heart of the discussion is the relationship between the “body” (robot) and the “brain” (AI) and their interconnection. The document expounds on the confine between (mechanical) automation and (human) autonomy, outlining possible scenarios arising from recent developments in robotics applied to different contexts, with the aim of avoiding confusion, excessive optimism or catastrophising and providing a balanced analysis of the roboethical and legal issues which will define the future governance of these new technologies.

Particular attention is paid to the replacement of human labour with robots and to new, specifically human jobs which cannot be replaced by technology; to dependence on robots (intended as social or personal dependence, or as “technological vulnerability”); to providing information to the public in order to boost critical awareness and encourage the “metabolisation” of innovation; to the robotic divide and inequality in access to technology, whether due to its cost or to a lack of the skill and motivation necessary to use it; and to the responsibility and liability of scientists and the ethical and professional codes of conduct of the designers, which refer to the principles of human dignity, privacy and safety.

The document also focuses on applications in medicine and healthcare (potential and limitations of robotic surgery, especially in experimentation; robotic assistance or robotics for assistance; biorobotics and neurorobotics); on the use of robots for military and surveillance purposes; and on possible new forms of legal liability.

The Committees conclude by offering various recommendations for society. These pertain to the need for critical awareness, the desirability of an interdisciplinary analysis of the social impact of robotics, the consideration of justice and non-discrimination, the need for ethical codes for robot programmers and for ethics committees for robotic research, and the importance of studying ethics in engineering and IT courses.

In medical contexts, the Committees affirm the need for a weighing-up of the risks, costs and benefits in robot experimentation and application and for equity of access; while in relation to military use, there is a need to incorporate the study of ethical issues in military robotics and an urgent need for the international community to take a position on the development of the infant technology of autonomous weapons.

Finally, from a legal perspective, the need for clarification of the new meaning and limitations of human legal liability in relation to robots, the need to protect public safety and the desire for European legislation are affirmed.

The working group was coordinated for the Biosafety Committee by Professors Andrea Lenzi, Carlo Caltagirone (who is also a member of the Bioethics Committee), Roberto Cingolani and Roberta Siliquini and for the

Bioethics Committee by Professors Lorenzo d'Avack, Laura Palazzani, Riccardo Di Segni, Salvatore Amato and Luisella Battaglia.

The opinion benefited from a hearing with Prof. Roberto Cingolani and from the observations and comments of Professors Antonio Amoroso, Carlo Casonato, Bruno Dallapiccola, Pier Angelo Morandini and Pier Franco Pignatti and of Dr Carlo Petrini.

The opinion makes use of figures provided during the hearing by Prof. Roberto Cingolani and made available for publication.

The document was approved unanimously by members of the Bioethics Committee present at the meeting of 26 May 2017 (Professors Salvatore Amato, Luisella Battaglia, Carlo Caltagirone, Stefano Canestrari, Carlo Casonato, Francesco D'Agostino, Bruno Dallapiccola, Antonio Da Re, Lorenzo d'Avack, Riccardo Di Segni, Carlo Flamigni, Paola Frati, Silvio Garattini, Marianna Gensabella, Assunta Morresi, Andrea Nicolussi, Laura Palazzani, Massimo Sargiacomo, Monica Toraldo Di Francia, and Grazia Zuffa; and the ex officio members, Drs Maurizio Benato and Carlo Petrini), and by members of the Biosafety Committee on 17 July 2017 (Professors Carlo Caltagirone, Paolo Gasparini, Maurizio Genuardi, Marco Gobbetti, Paola Grammatico, Piero Angelo Morandini, Luigi Naldini, Ferdinando Nicoletti, Giuseppe Novelli, Pier Franco Pignatti, Roberta Siliquini, and Paolo Visca).

Prof. Cinzia Caporale for the Bioethics Committee and Professors Antonio Amoroso, Antonio Bergamaschi, Roberto Cingolani and Mauro Magnani for the Biosafety Committee were absent for the vote but expressed their approval at a later date.

Background

This document, prepared by the working group of the Italian Committee for Bioethics in conjunction with the Italian Committee for Biosafety, Biotechnology and Life Sciences, reports the group's observations and opinion on robotics with the aim of helping to establish the Italian position, including in relation to European projects for the ethical and legal regulation of robotics[1].

I. Brief description of the development of robotics and its technological applications

1. Preliminary definition and distinctions: Robots and AI

Arriving at a precise definition of "robot" is no easy matter, given the rapid ongoing evolution of both the field of robotics and its significance in the relationship between ethics and technology (or in the consideration of technology from the perspective of ethics)[2]. The term "robot" was coined by K. Capek in 1920 as a replacement for the term "automaton", which indicated an artificial, mechanical machine built by humans to carry out precise operations, originally in the context of labour (in Czech, robot indicates work). However, robotics has made extraordinary progress in the last sixty years. From their past

---

[1] *See Report with recommendations to the Commission on Civil Law Rules on Robotics issued by the European Parliament Committee on Legal Affairs,* 2016*.*

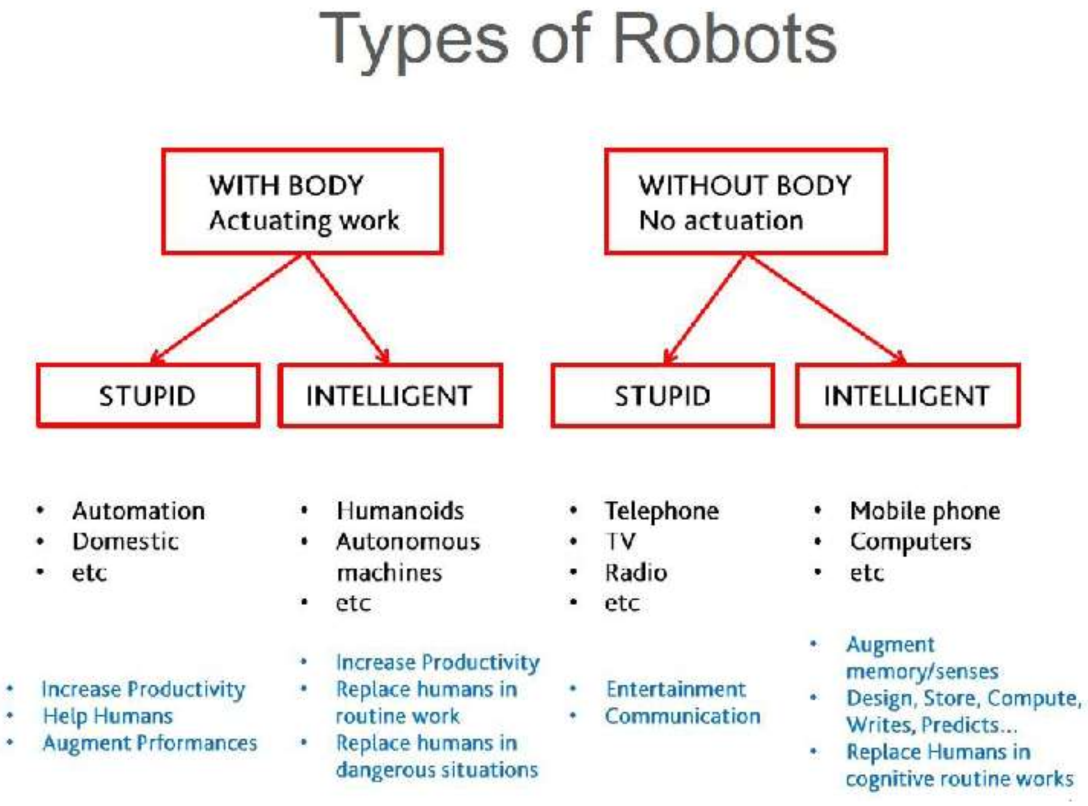
In Ancient Greece, the term *Banausia* (from *banausos,* meaning *artisan, manual work, craftsman*) was used to refer to manual labour and the mechanical arts in general. Its use had a negative connotation, as artisans and manual labourers were considered by the Ancient Greeks as an inferior class. Between 1400 and the early 18th century the value of such work was progressively re-evaluated, given that some of the procedures used by workers and artisans to modify nature are based on knowledge of the natural world.

The defence of the mechanical arts from accusations of unworthiness and the refusal to equate manual labour with slavery entailed the abandonment of a millennia-old image of science and the end of any distinction between *knowing* and *doing*. In this context, the great philosopher Francis Bacon was the first to tackle the fundamental ambiguity of the mechanical arts. Even 400 years later Bacon's intuition is still relevant, and indeed has acquired even greater weight in the era of robotics and artificial intelligence.

This ambiguity persists, albeit in a different way, in the transition from ancient to modern technology, as demonstrated by Heidegger (M. HEIDEGGER, *La questione della tecnica*, in *Saggi e discorsi*, editor G. VATTIMO, Mursia, Milan 1980). The quantitative increase in technological power has led to a qualitative leap: modern technology does not exploit nature merely to satisfy human needs, without disturbing its balance, but with its presumptuous mining and stockpiling disrupts the equilibrium between humans and nature. Nature is instrumentalised by humans, but only because humans, for their part, are already instrumentalised by technology. It is no longer technology that is a tool in the hands of humans, but humans who are its "functionary" (M. HEIDEGGER, *Perché i poeti*, in *Sentieri interrotti*, editor P. CHIODI, La Nuova Italia, Florence 1979). Goethe's image of the sorcerer's apprentice, who creates a mechanism he is unable to control, is an example of Heidegger's concept of a functionary of technology and the "danger" seeming to lurk in the essence of modern technology. However, this danger is not inevitable. Heidegger himself sees the ambiguity in the essence of technology, and it is through this very ambiguity that we can be freed from the restraints to which modern technology seems to bind us, from our dependence on its objects, from the urge to perfect them ever more, by behaving "differently". This "differently", which for Heidegger means doing without the gadgets of technology and the "giving up" of things, (M. HEIDEGGER, *L'abbandono*, editor C. ANGELINO, Il Melangolo, Genoa 1983), for Hans Jonas takes on the strong significance of a commitment to rethink "an ethics for the technological age" consistent with the imperative of responsibility (H. JONAS, *The Imperative of Responsibility*, University of Chicago Press, Chicago, 1985).

as static mechanical objects which received and executed specific repetitive orders, robots have become autonomous, mobile creations capable of carrying out general actions, learning, and adapting to their environment in a not always predictable way, even when not under human control; they are capable of analytical cognition, communication (word recognition), expression, and emotions - or, at least, the mimicking of human emotions in their outward appearance. And in the not so distant future, a certain scientific optimism looks forward to robots with artificial intelligence (AI) and thinking, decision-making and self-determination skills similar to that of humans. It is not surprising that the most usual representation of a robot, thanks in part to the mass media, literature, films and TV series, is of a mechanical "body" which thinks and behaves like a human being[3].

The first distinction to be made is between robots with and without a body (figure 1). Having a body means being able to generate movement, i.e. to produce physical work, unlike a computer, which is immobile and hence unable to perform any kinematic movements. Both robots with and without a body may be "stupid" or "intelligent", that is, with or without cognitive skills.



(Figure 1)

A "stupid" robot with a body generally carries out automated work (e.g. mechanical work, domestic work, etc.), increasing the productivity and performance of the humans controlling it, without making any independent decisions: whatever it does, whether right or wrong, depends on its program and operator, i.e. the human.

---

[3] Examples include *Blade Runner; AI; Matrix* and its sequels; *I, Robot*; *Ex Machina*, etc.

An "intelligent" robot with a body may be humanoid or an autonomous machine, not necessarily anthropomorphic. It too increases productivity, replacing humans in routine tasks or dangerous activities (e.g. working on a production line, carrying out activities in emergency situations, war zones or hazardous or radioactive areas, etc.)[4].

"Stupid" machines without bodies (TV, radio, telephone, etc.) used for entertainment and communication have long been a part of everyday life, to which humans are now "addicted".

Their "intelligent" counterparts (from smartphones to supercomputers, distributed or artificial intelligence, Google, Cloud computing, etc.) help humans to plan, to accumulate data and images, to make calculations and predictions and to write, replacing them in routine cognitive work.

A key point which must be clarified to enable the development of an ethical and legal position is the distinction between "body" (robot) and "brain" (AI), as well as their interconnection: the body and brain cannot be separated, because neither is dominant - it is the interconnection and orchestration between the two which is fundamental.

The history of robotics and AI seems to confirm that scientists and researchers are fixated on the creation of super-intelligent robots. The expression "artificial intelligence" was coined in 1956 by the American mathematician John McCarthy; to this day, scientists and philosophers are still debating the actual "intelligence" of machines. In most cases, the objective of AI researchers is to achieve what currently appears to be one of the most inaccessible scientific goals: to understand the principles and mechanisms of the function of the human brain in order to reproduce human intelligence in a machine. As far back as the 1960s published articles identified many of the obstacles which research into AI would encounter over the coming decades, essentially due to the huge difference in the results achieved in comparison with the cognitive functions of humans themselves. In 1969, *Perceptrons* by Marvin Minsky and Seymour Papert revealed the limitations of the first artificial neural networks, which the authors themselves had developed [5]. Given the authority of its authors, the publication of *Perceptrons* led to a significant reduction in scientific interest and investment, which endured until the early 1990s. Then, however, there was a rekindling of interest in AI, and research concentrated on the "intelligent agent" as an entity. This led to numerous studies of intelligent computer programs and intelligent agents "incorporated" into a physical system, such as computers with cognition[6].

The hope that humans could build intelligent humans was reignited, partly due to the new generation of computers[7]. Scientific progress in robotics tends

[4] The Istituto Italiano di Tecnologia in Genoa is currently building a centaur with a human-like upper part and an underpart which can move around on rock in places where vehicles with caterpillar tracks cannot reach, and can thus be considered a post-evolutionary machine.

[5] M. MINSKY, S. PAPERT, Perceptrons: An Introduction to Computational Geometry, The MIT Press, Cambridge MA, 1969.

[6] In 1997 IBM's Deep Blue computer, which could process 11 billion operations every second, created history by beating Garry Kasparov, the world champion, at chess. But it became clear that Deep Blue could not actually think, which dampened the enthusiasm for research into AI.

[7] After the second international conference on "Beneficial AI" held from 6-8 January 2015, the document "Asilomar AI Principles" was adopted. This contained 23 principles for the development of artificial intelligence. These ranged from research strategies and legal implications to the consequences for future generations. They are grouped into three areas: research; ethics and values; and longer term issues.

above all to provide systems to solve important problems for humans and society[8]. Once again, the question arises: must we already worry that human beings will soon be obsolete? As Minsky wrote: “Will robots inherit the earth? Yes, but they will be our children!”[9].

It should be noted that the question of significance and the ethical issues posed by current and forthcoming technological developments are not new to the history of human thought. The possibility of creating beings with human characteristics by alchemic, magical or mystical means has been contemplated in various cultures and expressed in myths and legends. These not only consider the creation of useful beings, but also the risk that such creatures might be uncontrollable and that a means to deactivate them is necessary. Today, all this might no longer be just a legend[10].

2. Recent documents

Debate on robotics is growing both in Europe and worldwide, although there are different views on the outlook. Each of the categories of robots described above has its own issues in relation to sustainability, positive and negative impact and ethics and regulatory problems, which will have to tackled ad hoc.

The most important recent documents in the field include *Artificial Intelligence, Automation, and the Economy* published by the White House’s Office For Science and Technology in October 2016, and the *Report with recommendations to the Commission on Civil Law Rules on Robotics* issued by the European Parliament Committee on Legal Affairs in May 2016 and transformed into a Resolution in February 2017.

The first (figure 2) focuses on artificial intelligence rather than robotics, and specifically on how a “good AI society" is organised. This document, which was written by Silicon Valley experts, is highly optimistic: AI should help improve everything, including cyber war and autonomous weapons. The report’s ethical recommendations are limited to ensuring that everything concerning machines is made transparent.

(Figure 2)

[8] P. MC CORDUCK, Machines Who Think, Natick, MA: A. K. Peters, Ltd., 2004 and S. RUSSELL, S. NORVIG, Artificial Intelligence: *a Modern Approach, New York, NY 2003; P. HUSBANDS, Robotics, in K. FRANKISH, W. M. RAMSEY (eds.), The Cambridge Handbook of Artificial Intelligence, Cambridge University Press, Cambridge 2014.*

[9] MINSKY, Ibid.

[10] The Jewish legend of the Golem is emblematic in this sense (see B. HENRY, *Dal Golem ai cyborgs. Trasmigrazioni nell’immaginario*, Belforte, Livorno 2013).

The second document (figure 3) specifically discusses robotics, i.e. *Good Robotics Society*, not *Good AI Society.* Good Robotics Society is based on the analysis of how many jobs would be lost with robotics and the need for “hard” and “soft” laws (i.e. how much impact they have and who is responsible for regulating violations), an agency for robotics and AI and a legal framework.

Figure 3

In April 2016 a study group from the Convention of the Society for the Study of AI and Simulation of Behaviour formulated five rules for the management of intelligent machines in the attempt to affirm that intelligence and robotics go together.

Rule no. 1: robots are multi-use tools; Rule no. 2: Humans, not robots, are responsible agents; Rule no. 3: Robots are products. They should be designed using processes which assure their safety and security; rule no. 4: Robots are manufactured artefacts. Their machine nature should be transparent; Rule no. 5: The person with legal responsibility for a robot should be attributed.

3. *Robots with a mechanical body and artificial intelligence: possibilities and limits*

The analysis herein considers machines with both a body and intelligence rather than machines without a body.

As already said, the latter are controlled at a distance by a human or programmed to perform given tasks.

Robots of this kind already exist in various sectors for various purposes, but without humans to make their decisions, they would be useless machines.

In contrast, scientists have been working for more than 50 years to develop “autonomous” robots with AI which are capable of thinking for themselves regardless of human input.

Their success will raise major safety, ethical and legal issues.

To avoid confusion, excessive optimism or catastrophising, it should therefore be remembered that the following observations refer to robots with both a mechanical body and AI.

In this sense, the biological body and intelligence of humans correspond to the mechanical body and artificial intelligence of humanoids while the brain-body coordination of humans corresponds to the AI-body coordination of humanoids.

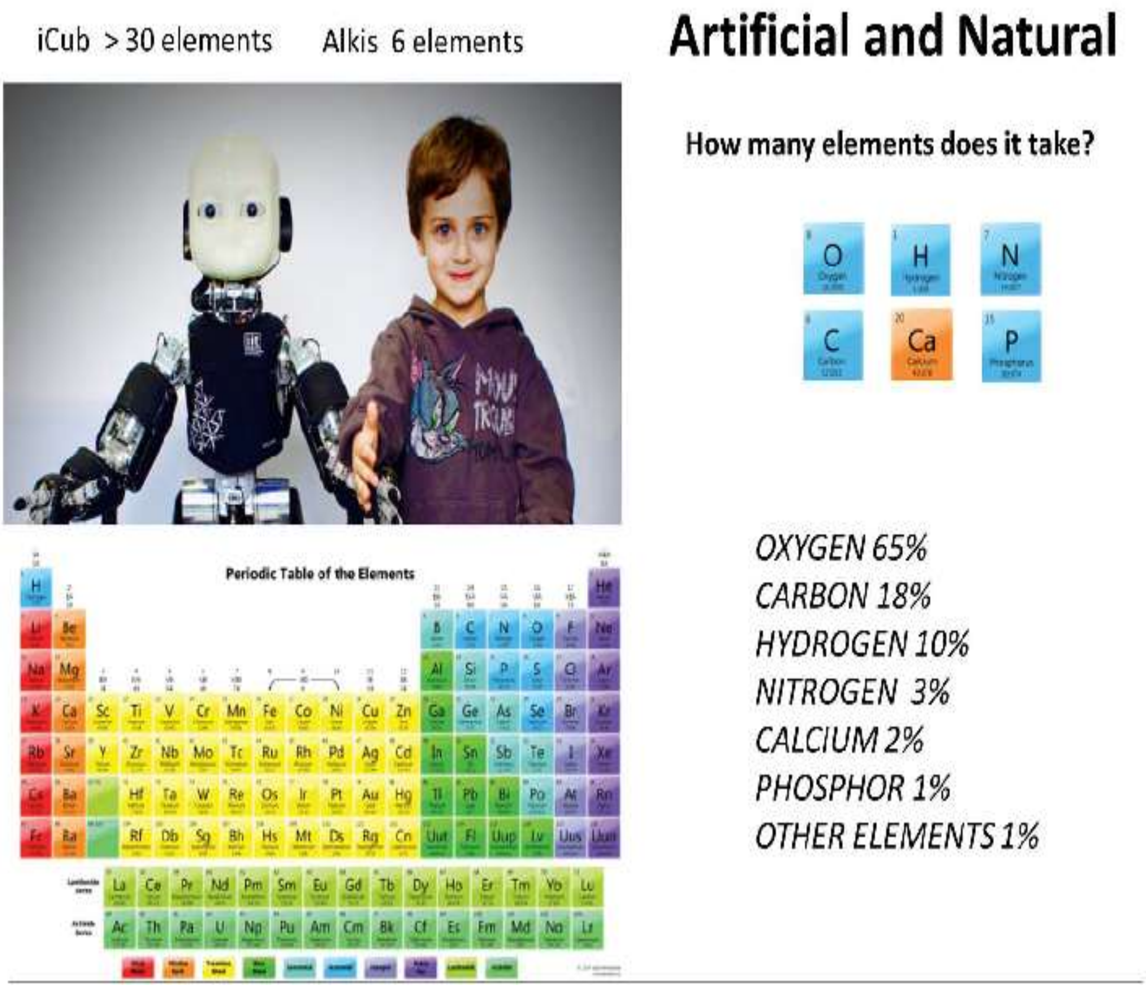
Any human-robot combination, or the possibility that humans and robots might one day be comparable, is currently unimaginable.

From the perspective of the mind-body relationship, humans have a system which has been optimised over the last 3 million years, enabling us to adapt, learn, feel, recognise, and manipulate, and giving us stability, dynamic equilibrium, and so on.

Technology can provide only a pale imitation: for example, robots can be given balance using gyroscopes (such as those in smartphones or aeroplanes), but these cannot compare with the human vestibular system.

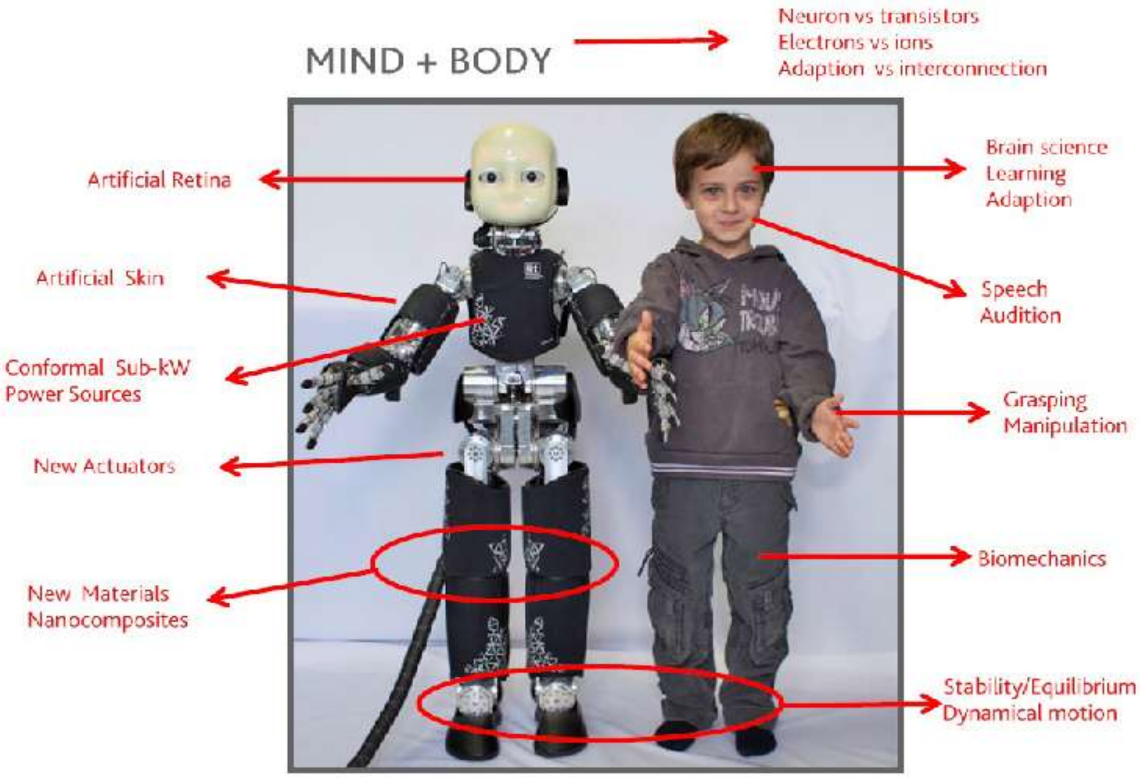
A human-robot combination would be even more problematic, given that 99% of the human body (and indeed of any natural animal or vegetable matter) is made from and grows using just 6 elements - oxygen, carbon, hydrogen, nitrogen, calcium and phosphor (figure 4) and is designed to break back down into the same substances at the end of life:

in contrast, any artificial machine requires from 30 to 50 different elements and is designed to be assembled as quickly as possible, without considering the need to disassemble it for recovery of its materials at the end of its lifespan (it takes 4 hours to build a car, but 40 to disassemble it, and in the meantime many of its parts will have deteriorated).



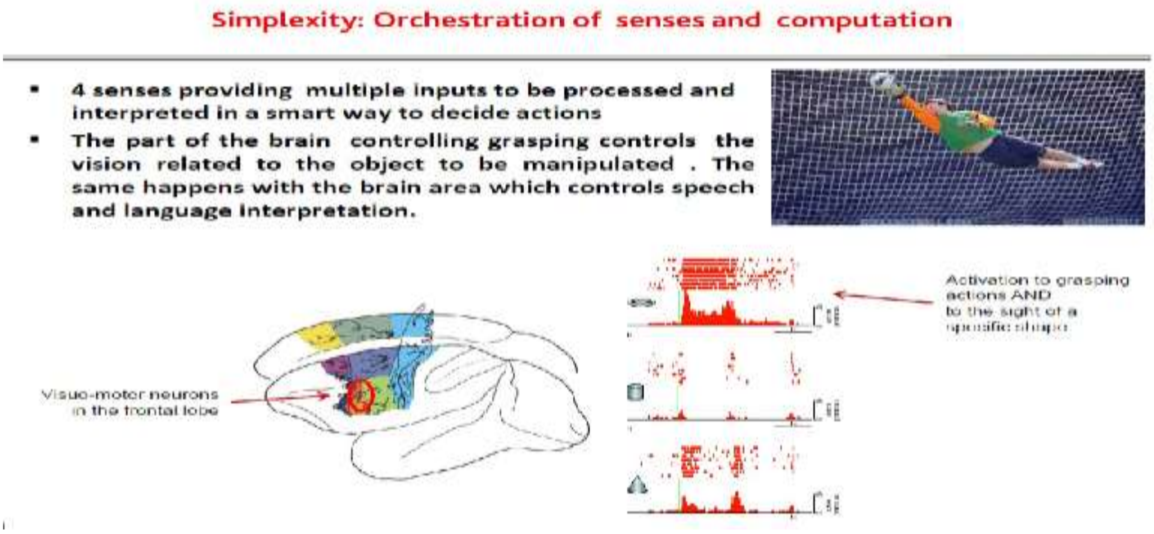
(Figure 4)

The mind-body connection will thus be difficult to equal with a simple union of a computer connected to movement actuators and sensors (figure 5).



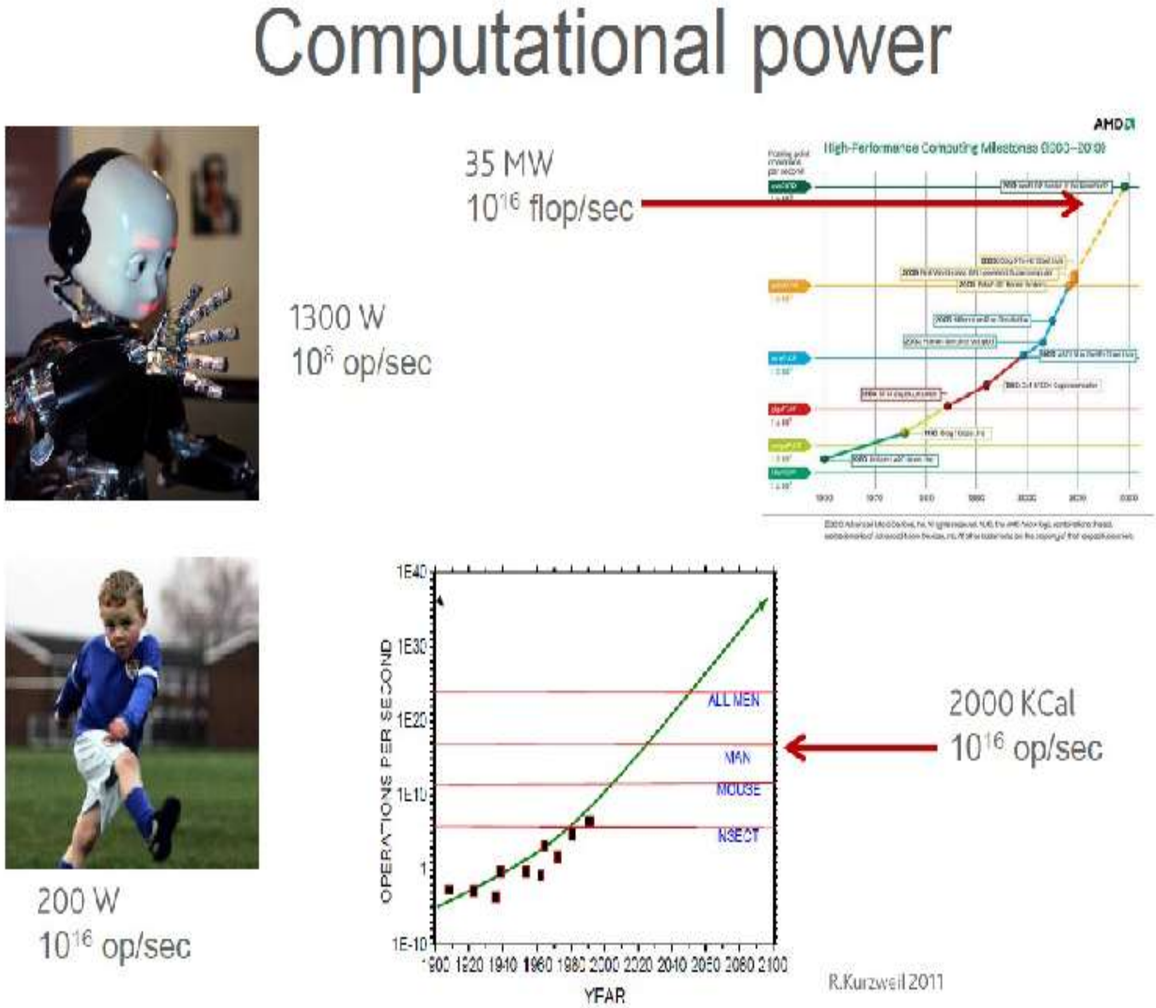
(Figure 5)

Furthermore, in living systems the brain is designed to think synergistically. For example, the group of neurons which controls vision also supervises manipulation of objects; the group which controls speech also handles language comprehension, and so on. These mind-actuation synergies are impossible in machines because they are made of non-synergic, mechanical materials and do not have any fibres which contract upon chemical stimulation (figure 6).



(Figure 6)

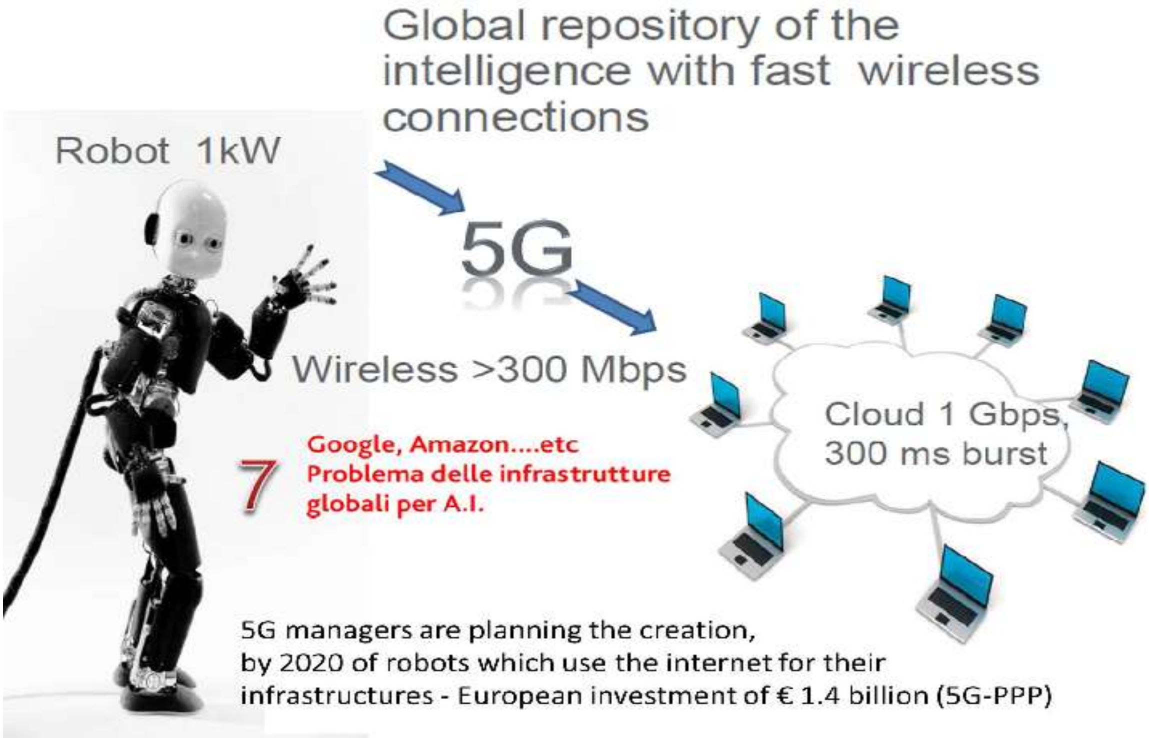
From a strictly theoretical perspective, it should be possible to increase the calculation capacity of machines (figure 7) to 10^{16}-10^{17} operations a second, equivalent to the human brain. Today's supercomputers can achieve such performances, but only by using 30 MW of power; moreover, they are the size of a room and required enormous cooling systems and an independent power generator.



(Figure 7)

A humanoid robot might therefore one day be able to think like a human, but it would have to have a power generator on its shoulders, its cost would currently be too high and it would be unusable. It is thus technically impossible to hypothesise a self-propelled system capable of thinking like a human and with the same mental as well as biomechanical properties.

Big Data companies which manage large computers and storage devices (figure 8), such as Google, are trying to find solutions to these problems: the robots will continue to be "stupid", with a small brain (10^8 operations per second) to enable their free movement, while the intelligent part will be in the Cloud, which will manage the apps used by the robots.



(Figure 8)

In any case, there are enormous difficulties to be overcome. For example, the ability to correctly interpret simple human gestures (body language) is extremely complicated to teach to a machine, as are speed of execution and the ability to adapt and understand intentions. Furthermore, these technologies will not be available everywhere, as they require internet, Wi-Fi and other infrastructures. In essence, companies which do not build robots are much more interested in this business than those which do, because the latter are well aware that robots alone will always be inferior to humans, while the former believe that a single worldwide artificial intelligence could offer high performance. The ethical issue thus exists because when sufficiently sophisticated robots and sufficiently powerful computers arrive, just a few companies will hold the entire global repository of intelligence in their hands. From a legal perspective, there is a need to avoid creating monopolies of businesses which could possess a daunting amount of highly sensitive and personal data as well as having the ability to regulate the behaviour of all the countless robots connected to their AI centres.

Having said all this, it is difficult to believe that one day there might be a robot as “intelligent” as humans, at least for as long as they are based on silicon, especially as robots were invented by humans to be used as “slaves”. What is needed is not a flesh and blood imitation but an intelligent machine able to do what it is designed to do. This requires the collaboration and convergence of numerous disciplines. Building a robot needs electronic and mechanical engineers, computer scientists, psychologists, neurologists, cognitive scientists, AI experts, logicians, mathematicians, philosophers, legal experts, economists, designers and artists all to work together.

II. Robotics, roboethics and society

1. Background

Interactions between humans and machines are expected to change quickly and ubiquitously once the latter are provided with the ability to reason or with sensomotor coordination. New prospects are opening for the use of robotic systems and AI as assistants to humans in various social (industrial labour, domestic work, information selection, problem solving) and medical settings. These inevitably give rise to new issues of applied ethics which, in the context of new forms of human-machine interaction, lead us to consider the "traditional" ethical principles (dignity, identity, safety and security, individual and collective responsibility, equal access to technological resources, freedom of research) in a different way.

2002 saw the birth of the new discipline of Roboethics[11], which deals with the ethical and social aspects of robotic technologies in their interaction with humans and human society as a whole. “Roboethics is an applied ethics whose objective is to develop scientific/cultural/technical tools that can be shared by different social groups and beliefs. These tools aim to promote and encourage the development of robotics for the advancement of human society and individuals, and to help preventing [sic] its misuse against humankind” (Veruggio, 2002).

Spyros G. Tzafestas offered a more recent definition of roboethics: “the branch of applied ethics that studies the positive and negative implications of robots to society, aiming at inspiring the moral design, development and use of robots, especially of intelligent and autonomous robots” [12].

Clearly, robotic science is still “in progress”: its products are progressively replacing a multitude of human activities and have a varying but in any case high social and ethical impact, given the many possible current and near (or distant) future applications. There is thus debate over how robotics can develop in a way compatible with humans and with respect for both individual and societal human dignity.

The robotic revolution will cause many changes in civil society and everyday life, affecting our play and recreation, our home, school and working life, our healthcare and transport, the organisation of our cities, our safety and security, the maintenance of public order, agriculture and energy production, and the protection of the environment, as well as our military defences. The future objective is to build autonomous living artefacts fit for various uses, albeit with different degrees of autonomy in their behaviour and task management.

2. Replacement of human labour and new jobs

To the extent that machines take menial, gruelling or dangerous jobs away from humans, the replacement of human with robotic labour is actually ethically desirable. In the same way, the use of mechanised labour to bridge gaps in human activities, such as in the assistance of vulnerable people, is also

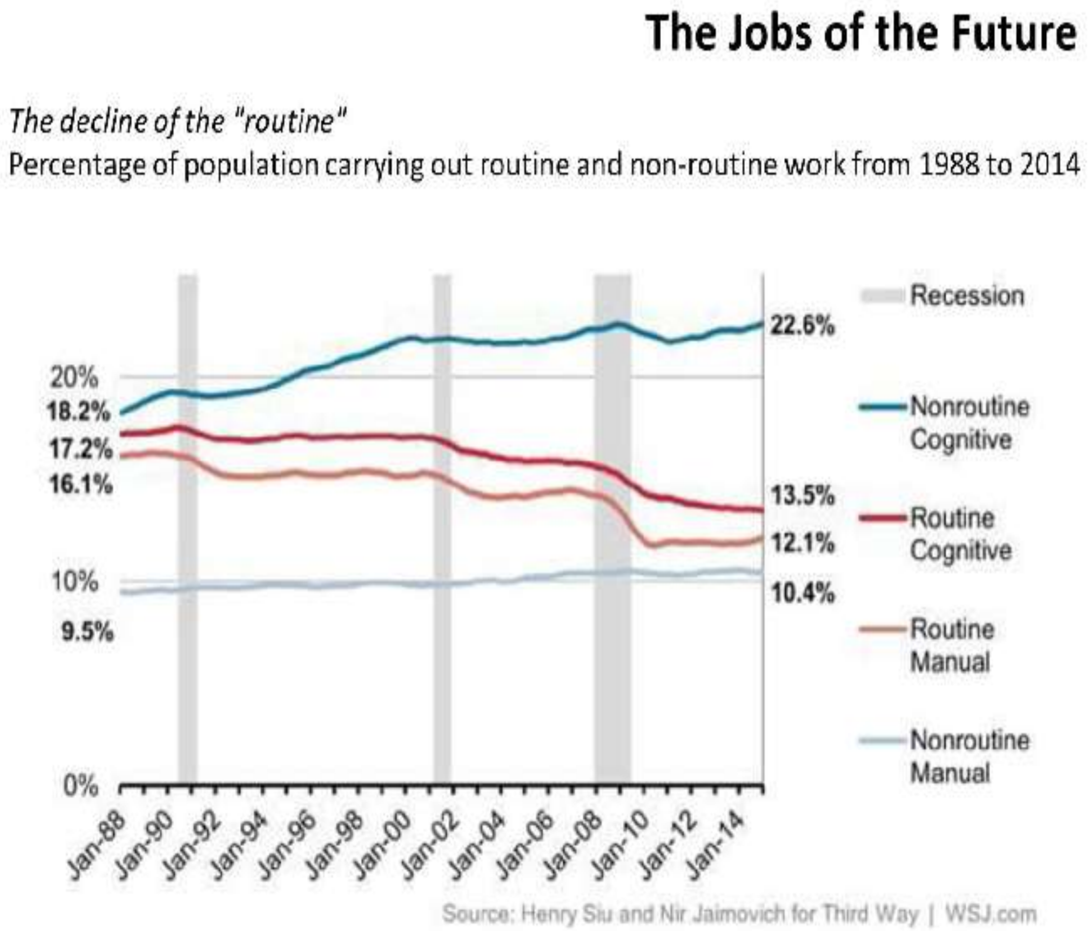
[11] Roboethics took shape in the first international symposium on this topic, organised by Veruggio in Sanremo in 2004. Even at this early stage, three different positions in relation to ethical responsibilities emerged in the robotics community.

[12] S. G. TZAFESTAS, *Roboethics. A Navigating Overview*, Springer, Dordrecht 2016.

desirable. The advent of home automation should enable us to live in “intelligent homes” where special sensors can actuate and monitor the operation of equipment, systems and home appliances.

However, with regard above all to labour, the robotic revolution will pose the problem of how to manage the progressive replacement of humans with intelligent machines which work more efficiently, quickly, precisely and cheaply. The introduction of robots to replace humans in today’s society could cause social tension (e.g. job loss[13]). Such situations will have to be managed by balancing the benefits (efficiency, economic savings, technological competitiveness in the international market) and the risks/harm (loss of human jobs, pension and benefit issues, greater socioeconomic inequality).

However, as can be seen from figure 9, over the last 25 years there has been a rise in non-routine cognitive jobs using the “brain" (i.e. skilled labour: craftwork, artistic work, etc.), a drop in routine cognitive jobs (which can be replaced with machines) and a moderate growth in non-routine skilled jobs (e.g. plumber).



(Figure 9)

[13] Bill Gates, the founder of Microsoft, in the ongoing worldwide debate on the rise of robots in factories leading to human job losses, proposed that robots be taxed to create an unemployment fund.

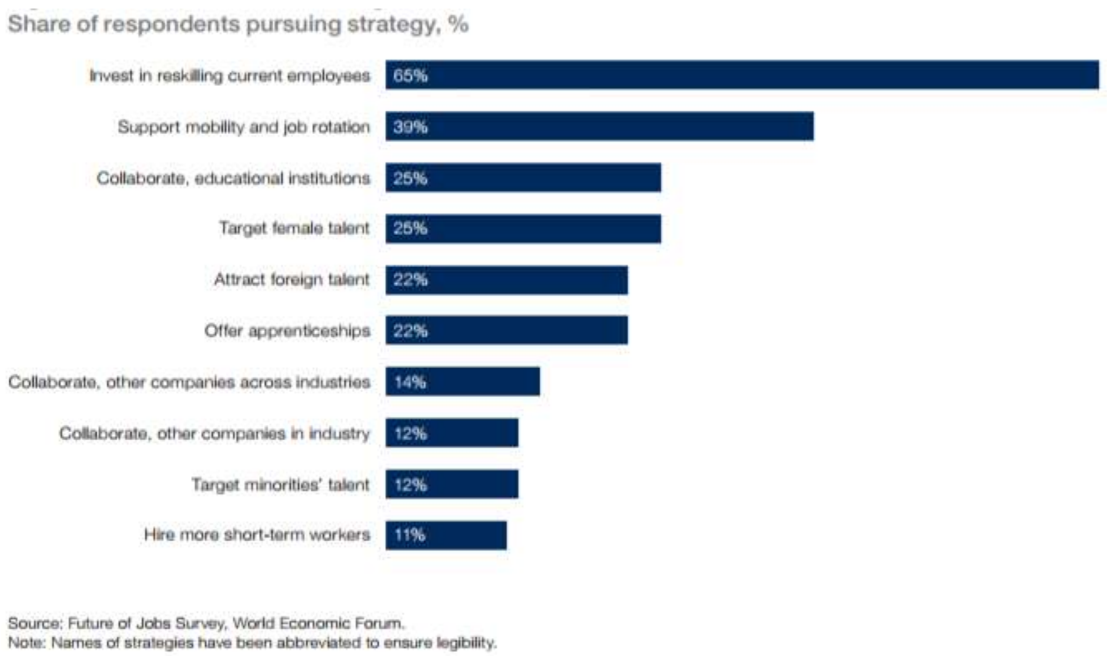
It can be deduced that non-routine work will never disappear because no robot is capable of performing it, as it will never be able to carry out the necessary categorisation.

The point is that humans are cognitive beings who recognise objects by their function (e.g. water glasses, wine glasses and tankards are all recipients for drinks), while a robot recognises them not by function, but by shape (if the only drink recipient it has memorised is a water glass, it will be unable to provide any other type).

Teaching a robot to reason by function rather than shape requires a cognitive approach which while currently possible, is highly complex and extremely costly, as already seen above in the differences between humans and machines.

The latest data from the World Economic Forum's survey on the future of labour (figure 10) revealed the need to invest in employee reskilling, i.e. training them to perform different roles.

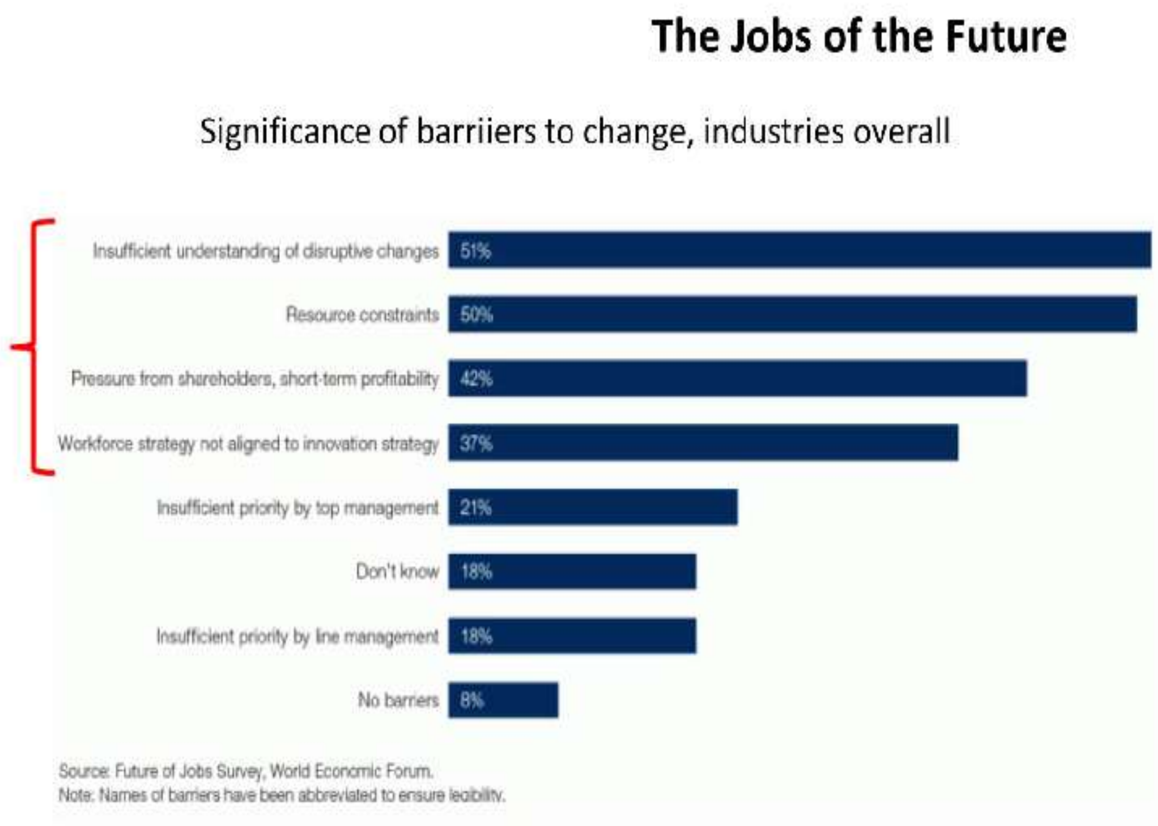
Future workforce strategies, industries overall



(Figure 10)

Society should look at this continuous learning process from an ethical perspective.

In the survey of barriers to change in an industrial context (figure 11), the problem is not only to successfully predict disruptive changes which cause industrial catastrophes but also to anticipate new businesses.



(Figure 11)

New job profiles are envisaged for the future (figure 12): nurses with computer experience; home care for the third and fourth age; body-part makers; doctors of nanomedicine; bioinformatic experts; bioarchitects; digital architects (cloud controllers); material architects (3D printing, recycling, sustainable materials, water cycle, waste); energy managers; food technologists (traceability, analysis, packaging).

(Figure 12)

3. *Dependence*

Social dependence on robots is generally intended to mean dependence on machines; the evolution of robots in the short term is likely to be comparable to the computer revolution. In just a few years we became dependent on computers (mobile telecommunications, smartphones, computers) - dependence on robots is a likely scenario in the near future, as they work their way into our everyday life. However, a boost in robotic technology could increase human vulnerability. The issue of dependence on technology, intended as a personal or social dependence, will thus arise in a new form.

The difficulty, typically seen in individuals who live with the aid of educational, recreational, artistic and medical robots and intelligent systems, of distinguishing between what is real or natural and what is imaginary or artificial may lead to forms of individual dependence on robot technology. While the distinction between robot and human will probably remain clear to future users, if robots too closely resemble humans the interaction could arouse feelings, attachment, and dependency, especially in particularly vulnerable groups such as the elderly, people with disabilities, children, and people with special educational needs. Studies have been conducted to analyse the emotive impact of the aesthetics of robot design in relation to age, culture and individual personality. The options are to develop "opt out" mechanisms to block the robot before the process leading an individual to dependence is triggered (similar to alarm systems when there is excessive exposure to certain technologies), or to limit the humanoid features of robots to avoid boosting the affective aspect in addition to the functional aspect.

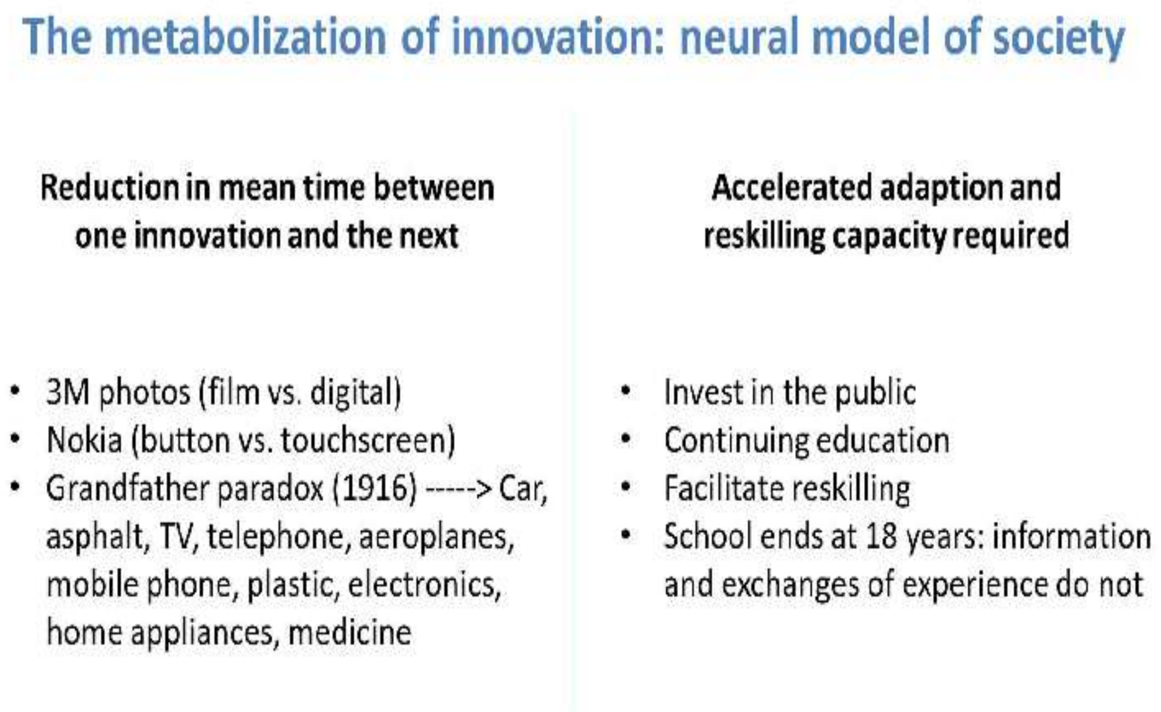
There is also the problem of "robotic deception": robots which act like humans imitate their behaviour, and for the unaware, this comprises a form of deception and delusion, which could be harmful.

4. *Information for the public and "metabolisation" of innovation*

It is essential to provide the public with incisive information. A significant responsibility in relation to the legitimate development of robotics falls initially on the scientists and technologists, who should increase public awareness of the social and ethical issues of robotics so that society may take an active part in the creation of a collectiveness awareness which can identify and prevent the improper use of modern technology. Users must be adequately informed of the opportunities and limitations of technology, partly with an eye to inclusion and participation in the definition of public and legislative policies.

This is particularly important to avoid, on the one hand, utopian hopes, and on the other, irrational fears which could divert attention from the real problems and, in the final analysis, produce illusory enthusiasm or a general unquestioning refusal of a technology which could be a tool for economic development and social progress, to the extent that it assists rather than replaces humans.

Society can be considered as an interconnected neural system which, to be sustainable, must give citizens time to metabolise innovation, which itself cannot be held back. The speed at which citizens readapt and are reskilled must thus be accelerated (figure 13), with citizens under constant training. Furthermore, the information they are given must be as precise and objective as possible, enabling them to reskill as best they can, should they lose their job.



(Figure 13)

There is thus a need to invest in public information (public awareness of technology) to facilitate, at all times, reskilling and adaptation to rapidly changing technology. The cases of 3M and Nokia are emblematic: the radical change in the core business of these companies over just a few years caused economic crises and huge job losses in their sectors and countries[14]. The ethics of an advanced society must explain and pre-empt these dynamics before they occur.

The time lag between technological acceleration and slow mental and social assimilation opens a gap which is hard to bridge. It requires continuous information and training in “flexibility” and “mobility” to try to minimise discrimination between the included (people who will be part of technological and robotic society) and the excluded (those who are unable to acquire the necessary skills over time).

5. Inequality: the “robotic divide”

To the extent that robots can provide humans with an efficient tool to perform tasks and improve quality of life, a socioeconomic assessment is ethically important to avoid the so-called robotic divide, a cause of inequality in access to technologies caused by their costs and by the skills and motivation to use them, acquired through a process of continuing information and training.

It should be borne in mind that the economic differences between different countries are fundamental in robotics as in other sectors, given that energy and water have a primary role for the functioning of a society which makes use of

[14] With the transition from push button to touchscreen technology Nokia, at the time the biggest mobile phone company in the world, collapsed and Finland’s GDP changed almost overnight. In a similar way, the rise of digital cameras has led to the disappearance of cameras using silver nitrate film, with a heavy impact on the core business of the world leader in photography.

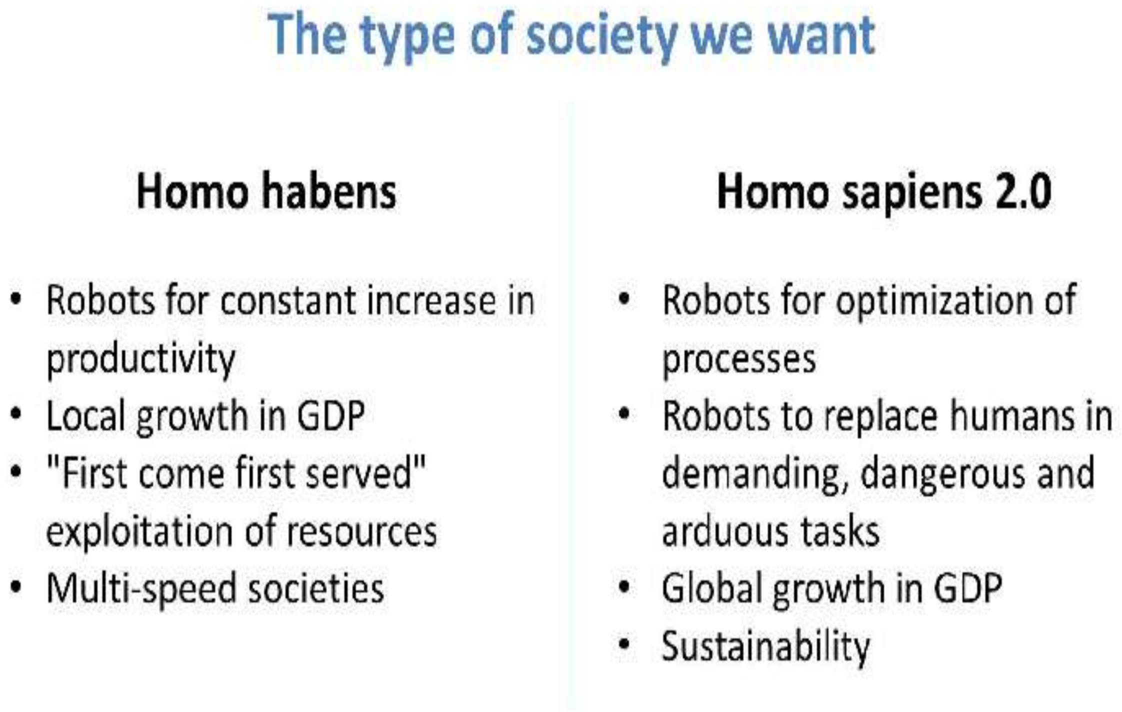
robotic machines. This is a worldwide inequality which translates into a global, or more accurately “glocal”, ethical problem: while global, it mainly affects G10 or G20 countries.

80% of energy is produced for the USA, Europe and Japan, which account for just 20% of the world population: all the rest are in the dark. Where there is energy there is water, so whoever has electricity also has water, industry, welfare, agriculture and a food production chain, all of which lead to a greater life expectancy. This is therefore a glocal problem, and a complete redistribution of resources is essential.

6. *Planning the social and ethical use of robots: the responsibility of science*

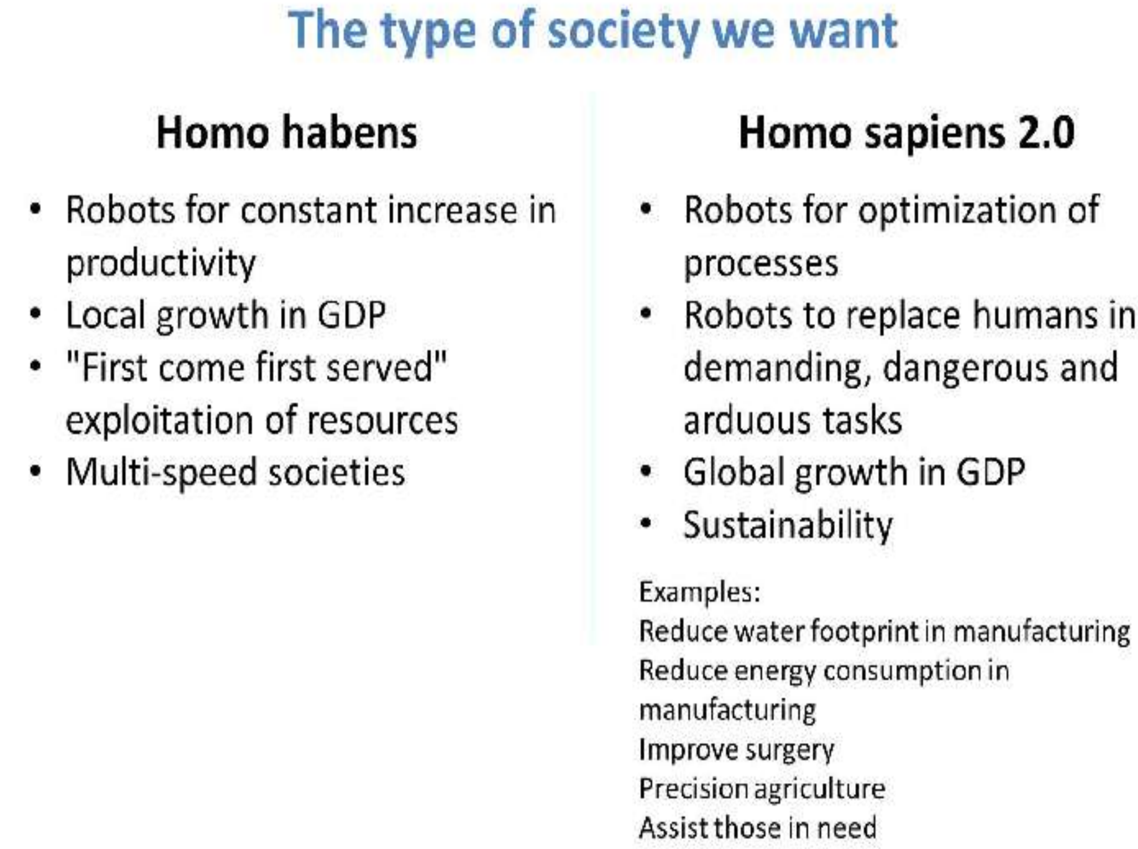
All these fields of robot application presuppose adequate design and programming; it is imperative that these activities must not be arbitrary or unrestricted, dictated by the vagaries of politics and the market.

Above all, it is necessary to understand what model of society we would like to see in 2100 (figure 14). Today’s model is *homo habens* (with robots to constantly boost productivity, growth of GDP on a local basis, exploitation of resources such as lithium, petroleum, etc.).



(Figure 14)

The future could see two models: one post- or trans-humanism, i.e. *homo sapiens 2.0* (like industry 4.0), which improves its performance so that robots - that is, machines to boost sustainability - optimise processes and replace humans in dangerous, demanding and arduous tasks, enabling GDP to grow on a global rather than local basis. In this ethical model of intelligent robotics (figure 15), robots are used not to increase GDP indefinitely but to reduce the water and carbon footprint of industry as well as its energy costs[15].



(Figure 15)

7. An ethical code for designers

Rather than the creation of a robot incorporating a prefixed ethical code, it seems more realistic to recommend and proscribe safety principles and parameters for the protection of future users to be implemented in the design and programming phase by IT experts and robotic engineers.

According to the European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, the guiding framework for the design, manufacture and use of robots should be based above all on the principles of autonomy, beneficence, non-maleficence and justice.

The principle of autonomy refers to the capacity to make an informed, un-coerced decision about the terms of interaction with robots. Beneficence means that robots must act in the best interests of humans. Non-maleficence means

[15] One of the main challenges is agriculture, where use of robotics for the targeted application of nutrients, herbicides and other pesticides could reduce environmental and agronomic damage and production costs. Robots controlled by GPS and able to reconstruct images, and thus intelligent enough to recognise diseased or invasive plants, could identify and treat solely crop plants lacking nutrients or under attack by parasites, or solely invasive plants, hence reducing the amount of substances used. This would be an example of artificial intelligence in a robot with a body, used to mitigate the impact of a production process.

that robots must not harm any human. The principle of justice requires the fair distribution of the benefits and economic accessibility of robots for home use, especially those for use in healthcare.

The principles enshrined in article 2 of the Treaty on European Union and in the Charter of Fundamental Rights of the European Union (including respect for human dignity, equality, non-discrimination, informed consent, respect for private and family life, protection of data, non-stigmatisation, transparency, and individual and social responsibility) and in the existing ethical codes must also be respected.

This Resolution includes a series of recommendations in relation to the registration of robots (to ensure traceability), civil liability (so that there is no limit to the damages which can be claimed and to ensure the establishment of an obligatory insurance regime) and the interoperability of robots connected via the internet.

A voluntary code of ethical and professional conduct for robotic engineers is also proposed, inspired by a number of general principles and guidelines for the actions taken by all interested parties. This code is intended to cover all research and development activities in the robotics sector, requiring the close cooperation of all disciplines so as to ensure that research into robotics in the European Union is conducted safely, ethically and effectively. All researchers and designers are thus required to act responsibly and in full awareness of the need to respect human dignity, privacy and safety. The Code also affirms the duty to respect fundamental rights, the principle of precaution, and the principle of inclusion (so as to ensure transparency and respect for the legitimate right of access to information by all interested parties). It recalls the duties of accountability (to take account of the social, environmental and health-related impact of robotics for current and future generations), reversibility (involving the ability to cancel the last action or sequence of actions, enabling users to cancel unwanted actions and return to the “correct” phase of their work), the protection of privacy, maximisation of benefits and minimisation of risks.

From this perspective, it is evident that the ethical review process should be conducted by sufficiently qualified people and be independent of the research itself, to avoid any conflict of interests between the researchers and the reviewers and between the latter and the organisational structures. For this reason the Resolution also proposes that a Research Ethics Committee (REC) of members with vast experience in the robotics sector be established within the organisation itself. This multidisciplinary committee should have an adequate balance of scientific, philosophical, legal and ethical skills and also include representatives from other sectors (healthcare, education, social services).

Science and ethics are closely connected and must not be separated. This is essential to avoid returning to a dichotomy that was overcome decades ago, both theoretically and in practice, by ethics committees in research institutes and health facilities worldwide, including Italy.

The committee’s role will be to scrutinise all research activities conducted by the organisation concerned and involving humans; to assure the independence, professionalism and timeliness of the ethical review; to safeguard the dignity, rights and welfare of subjects participating in the research; to bear in mind the safety of the researchers and the legitimate interests of other parties involved; and to formulate informed opinions on the scientific merit of proposals and informed recommendations for researchers in the event that their proposal should prove inadequate in some way. Monitoring of approved research activities until their conclusion is also envisaged to ensure their constant control, especially if any variations in the projects are anticipated. Monitoring must be proportionate to the nature and associated risk so as to arrive at an exhaustive ethical assessment. Where a study is thought to have been conducted in an unethical manner, approval may be revoked and its interruption required. A key concern is the protection of users, who should be able to make use of robots without any risk or fear of physical or psychological injury, based on the principle that the robot performs the tasks it was built for, albeit in the awareness of its perceptive, cognitive and behavioural limitations. At the same time, respect for the emotive needs and the physical and psychological fragility of humans requires that monitoring also takes account of the right to privacy, meaning that personal information may not be collected or used without the consent of the person concerned. On the other side, the duties of users include not to use a robot in any way which violates ethical principles and standards or to modify it for use as a weapon.

Scientists must receive critical training in the ethical issues of robotics through the introduction of courses on ethics in scientific faculties, hospitals, healthcare facilities and computer and technology businesses.

Within the development of specific principles of professional conduct in this sector, it is important that the various design, production, activation and management steps can be clearly identified and hence legally traced.

Manufacturers of robot technologies should be required to have the "burden of proof": that is, to demonstrate and justify the relevance and safety, in relation to potential risks and benefits, of the new technology they intend to develop and market. In this sense, the principles of the "integrity of research" should be ensured.

III. Biorobotics and roboethics in medicine and biomedicine

In the general context of the ethical aspects of the social applications of robotics, this document specifically address issues related to healthcare. Robotic applications in this sector are very new, and have an interdisciplinary nature covering medicine, neuroscience, engineering, nanotechnology and various humanistic areas (philosophy, ethics, sociology, psychology, etc.). Their very useful results are another reason to keep an open mind and consider what future these technologies might make possible.

The production of robots to assist humans in clinical diagnostics, surgery, rehabilitation, personal assistance and health monitoring (medical robots) is being trialled, offering new prospects for biorobotics and neurorobotics.

The various categories in this field are discussed below.

1. Robotic surgery

Robotics enables surgeons to work on the human body at a distance, using software to control the movement of a mechanical arm. This specific area of telerobotics enables surgical procedures to be conducted with greater precision and less invasiveness (minimal trauma, reduced procedure time and patient recovery time) than can be achieved by human surgery.

This application will have to develop alongside the possibilities offered by multiple communication channels, which send remote digital commands to the robot at the speed of light (ISDN and ATM networks), further developing the frontiers of telerobotic surgery, which are currently still limited. Telesurgery will offer the benefit of uniting multiple doctors in an integrated communication and patient management network, in which the operators, from their individual workstations, will be able to communicate and discuss their decisions together and act through the robot. Telerobotics involves a further transformation of the patient's relationship with the care system. Its effect on the patient-doctor relationship will not be neutral: patients must be guaranteed not only the deftness and precision offered by the combination of a highly reliable robot and a technically competent surgeon but also the respect of all those virtues which accompany the fulfilment of a professional obligation, and only after clinical, ethical and professional evaluation.

Robotic surgery has already been introduced into healthcare in some sectors (general surgery, urology, gynaecology, orthopaedic, heart surgery), and some ethical requisites are necessary for its application.

First, its safety and efficacy must be tested. It is essential to prove the technique's reliability with randomised controlled trials demonstrating its safety and scientific efficacy, its real benefits and its potential risks or errors in comparison with conventional surgery (reduced time, invasiveness, trauma).

In urology, as in other sectors, randomised controlled trials (RCTs) against open surgery have been conducted. However, none of these showed any significant benefits. Despite this, robotic surgery has become standard in some urological procedures, probably due to the advantages perceived by the surgeon, which even randomised trials are not always able to reveal[16]. It would be difficult to conduct another RCT in the area of radical prostatectomy, for example, because surgeons who have access to a robot use it for that procedure, and media pressure on patients would make them unlikely to agree to take part in a randomised trial[17].In other fields such as heart surgery, following initial enthusiasm the use of robotics is still being evaluated[18].

The trial and application of robotic surgery must in any case involve consideration of the following aspects: assessment of the clinical risks and benefits with adequate evaluation of the overall usefulness of such systems[19], their true economic and financial sustainability[20] and equality of access;

[16] D. STEFFENS, R. THANIGASALAM, S. LESLIE, B. MANECK, J.M. YOUNG, M. SOLOMON, Robotic Surgery in *Uro-oncology: A Systematic Review and Meta-analysis of Randomized Controlled Trials*. Urology. 2017 Mar 20. pii: S0090-4295(17)30266-2. doi: 10.1016/j., and, to a lesser extent, (S.K. SON, N.R. LEE, S.H. KANG, S.H. LEE, *Safety and Effectiveness of Robot-Assisted Versus Open Radical Cystectomy for Bladder Cancer: A Systematic Review and Meta-Analysis. J Laparoendosc Adv Surg Tech A*, 2017 Mar 28. doi: 10.1089/lap.2016.0437.

[17] J.W. YAXLEY, G.D. COUGHLIN, S.K. CHAMBERS, S. OCCHIPINTI, H. SAMARATUNGA, L. ZAJDLEWICZ, N. DUNGLISON, R. CARTYER, S. WILLIAMS, D.J. PAYTON, J. PERRY-KEENE, M.F. LAVIN, R.A, GARDINER, *Robot-assisted Laparoscopic Prostatectomy Versus Open Radicalretropublic Prostatectomy: Early Outcomes from a Randomised Controlled Phase 3 Study*, in "Lancet", 2016, Sep. 10, 388 (10049), pp. 1957-1066.

[18] M. PETTINARI, E. NAVARRA, P. NOIRHOMME, H. GUTERMANN, *The State of Robotic Cardiac Surgery in Europe*. Ann Cardiothorac Surg. 2017 Jan; 6(1):1-8. doi: 10.21037/acs.2017.01.02.

[19] M. RITROVATO, F.C. FAGGIANO, G. TEDESCO, P. DERRICO, *Decision-oriented Health Technology Assessment: One Step Forward in Supporting the Decision-making Process in Hospitals,* in "Value in Health", Jun 1 2015 18, 4, p. 505-511-7, p. 1097.

[20] G. TEDESCO, F.C. FAGGIANO, E. LEO, P. DERRICO, M. RITROVATO, *A Comparative Cost Analysis of Robotic-assisted Surgery Versus Laparoscopic Surgery and Open Surgery: the Necessity of Investing Knowledgeably*, in "Surgical Endoscopy and Other Interventional Techniques", 2016 Mar 16, p. 1-8.

adequate patient information (informed consent); establishment of liability in the event of malpractice (of the surgeon, designer or manufacturer); and the training of doctors in the use of robotics in surgery, without losing the habit of operating without robots. Given the current trend towards super-specialisation, robotic surgeons should be envisaged as a new position alongside traditional surgeons carrying out procedures which do not benefit from this technology. Robotic surgery is, and must remain, a means and not an end. The surgical robot is merely an aid for the surgeon, not a replacement.

2. Robotic assistance or robotics for assistance

The use of robotics to assist children, the elderly, people with disabilities (non-autonomous, absence of movement, non-use of limbs) and adult patients is already standard practice in some countries (Japan, Korea). Robotic assistance can take various forms, such as distance monitoring and control of health (telerobotics, through video cameras or GPS systems); rehabilitation assistance; assistance in activities of daily living (eating, drinking, dressing, moving around, etc.), and providing company (in cases of isolation or depression). Various terms are used, often in combination, according to the function: “health/assistive robot”, “socialised robots/socially assistive robots”, “service robots”, “carerobots/carebots”, “robotic nurses” or “nursebots”. One example already in use is the CareBot[21].

From an ethical perspective it is essential to specify the conditions for the legitimate use of robotics in this context. These comprise respect for safety, guarantee of the quality of the assistance and prevention of injury (prevention of the risk of erroneous mechanical actions by the robot which could be hazardous to health or quality of life), guarantee of autonomy and consideration of specific individual needs and preferences (informed consent), the proportionality of the assistance and adjustment to the use of robotics. The risks and benefits must also be weighed up, with specific reference to the human need for help and socialisation while limiting and monitoring invasion of privacy, and all while ensuring equality in the distribution of resources for investment in research and access to technology.

An evaluation of the conditions for the legitimacy of the use of robots must be performed on a case-by-case basis. This must first assess the potential impact of the robot’s use on the individual and on the welfare services; alternatives to robotic assistance should also be offered, and extra care should be taken in the case of people with cognitive difficulties.

It is essential to avoid robots replacing human relationships. The artificial “care” offered by a machine leads to the dehumanisation of that care and the objectification of the patient, who is perceived as a problem or burden requiring a technological solution. Caregiving is an intrinsically moral and irreplaceable human practice. It enables the development of human virtues and skills, within the empathy and reciprocity of an interpersonal relationship which enables an adequate approach to people in a particularly vulnerable condition. While carebots can free caregivers from repetitive and draining tasks (thus offering a potential benefit by enabling a more profound provision of human care), it is one-sided, lacking reciprocity and empathy. Robotic assistance can support

[21] S. VALLOR, *Carebots and Caregivers: Sustaining the Ethical Ideal of Care in the Twenty-First Century*, in “Philos. Technolog.”, Springer, 2011, 24, pp. 251-268.

human care and should be offered where there are no alternative solutions (due to a lack of human resources), but it should always be borne in mind that carebots can lead to isolation of patients and deprivation of their physical and psychological autonomy.

There is thus a clear need for the adequate training of health and social workers on the opportunities and limitations of robotic assistance.

3. Biorobotics and neurorobotics

Contemporary robotics has a double link with the development of neuroscience and cognitive science. On the one hand, improved understanding of the neural and cognitive aspects underlying the behaviour of living systems has often stimulated the development of efficient robotic systems capable of operating in relatively unstructured environments. On the other, in many cases the construction of robots has provided significant contributions to the progress of neuroscience and cognitive science, such as seen with neural networks. A relatively recent example is the development of robotics in neuromotor rehabilitation.

Biorobots are hybrid bionic prostheses which can be connected directly to the human body and are perceived by the brain as part of that body. For example, artificial limbs interfacing with the peripheral nervous system are designed with the ultimate aim of using both afferent nerve signals for motor control and afferent sensorial stimuli to restore the subject's sensitivity and correct any motor control errors. This research aims to restore lost physical functions.

The rehabilitation and recovery of upper and lower limb movement (through the use of exoskeletons controlled electronically through weightlifting and treadmill systems) is currently of great interest, and various studies have shown that robot-assisted therapy of the upper limbs, whether in the acute, subacute or chronic stage, leads to considerably greater functional improvement than achieved by conventional therapy.

The field of neurorobotics is looking to the reproduction of artificial models of the human brain, visual perception through sensors or artificial vision, and communication (including in non-verbal forms) between humans and artificial systems, including the generation and understanding of particular emotive states (affective computing). For example, biomimetics studies the biological processes of nature to improve technology, producing integrated artefacts in our bodies and brains (known as "bio-inspired" artefacts); these artificial limbs interface with the peripheral nervous system with the aim of reactivating and correcting motor control, thus restoring the subject's sensitivity. Hybrid bionic systems - robotic artefacts which can be connected directly to the human body and perceived by the brain as part of that body (living artefacts) - are also under study.

Robotic and biomechatronic (biological-mechanical-electronic) prostheses raise questions in relation to human identity, integrity and freedom. There are organs that cannot be replaced by a biorobotic transplant without losing this identity. Doctors indubitably have the task of assessing, in good conscience, "morphological freedom", meaning the legitimacy of an individual's request to modify their body with robotic insertions as they please, as an expression of their autonomy and liberty. There is also a need to consider the difference, which cannot always be defined, between therapy and enhancement, namely to what point robotics, biorobotics and neurorobotics are mere treatments and at what point they actually start enhancing human abilities.

Another case is the brain-computer interface (BCI), a technology which enables direct communication between neuronal activity and an external device. It is essentially based on the ability to read neuronal activity, process signals and send commands to the outside world. Many important objectives have been achieved in recent years which enable people with sever motor disabilities to carry out everyday but nonetheless complex actions, such as eating or drinking, by controlling a robotic arm through the power of thought. This is achieved through an invasive interface involving the surgical implantation of a sensor in direct contact with the brain (and other interesting new approaches are under development). The invasive method is necessary for complex actions like moving an arm (many degrees of freedom), but BCI technology has also evolved in the direction of partially invasive and non-invasive sensors, making it possible to control an external robot or turn on a light in a relatively simple way.

If it is true that biological inspiration can help in the development of more effective robots, is it equally true that the development of bio-inspired systems can contribute significantly to neuroscientific research? What can observing a bio-inspired robot teach us about the neurophysiological mechanisms underlying the production of adaptive and intelligent behaviours? What are the experimental limits and potentials of robotic simulations of neuroscientific theories? We are not ready to answer questions of this kind. As has already happened in numerous circumstances, the rapid progress of a scientific sector or technology exposes situations whose ethical issues have hardly been touched upon.

For example, brain implants raise numerous ethical problems. The surgical grafting of sensors and microchips in the brain generally produces clear, more experimentally useful, more precise and more reliable signals, but gives rise to medical concerns in relation to the risk of rejection and infection. On the basis of current scientific knowledge, only extremely limited use of implantable microchips can be considered permissible, and only for the purpose of protecting the health of the individual concerned, to the extent that there are demonstrated medical indications, that there are no less risky or less invasive therapeutic alternatives, and that the patient has been adequately consulted through an informed consent process. This would enable invasiveness to be minimised in relation to the aim of the procedure, thus ensuring its proportionality. Caution and precaution must be the dominant principles in this sector, given the scientific uncertainty in its applications. The use of such microchips for cognitive enhancement raises significant ethical concerns, given the invasiveness and risk. It is not morally justifiable to risk life and health to achieve a cognitive improvement (especially as there is no guarantee of success). There is also the issue of privacy[22].

Further research into the long term social, cultural and health-related impact of the various types of connection between robotics and neuroscience must bear in mind the necessary distinctions between active/passive, invasive/non-invasive and reversible/irreversible implants. Research cannot ignore the principle of precaution, which must be applied whenever there is a high but uncertain risk, and must take account of the fundamental principles of the

[22] European Group on Ethics in Science and New Technologies, Opinion n. 20 - 16/03/2005 - *Ethical aspects of ICT Implants in the Human Body*.

respect of human dignity, physical and psychological integrity, autonomy, non-discrimination, privacy and the right to identity and justice.

IV. Robotics in the military, policing and surveillance

Ethics in relation to military robotics cannot neglect the basic principle, as declared in the Italian constitution, that war is to be repudiated “as a means of attacking the freedom of other people and as a means of international dispute resolution” or the related observations already expressed by the Italian Committee for Bioethics in 2013 in its opinion on *Human rights, medical ethics and enhancement technologies in military contexts.* The problem of enhancement was thoroughly discussed in that opinion and its conclusions should also be considered valid herein. However, in this case the topic is not enhanced soldiers with exceptional abilities, but machines to replace soldiers.

The military application of robotic technology is a major area of development, in which prevention of the use of robots against human beings is a particular challenge. As far back as the second world war, the Germans and Russians had already developed wire- or radio-guided tanks. Various armies currently use crewless land vehicles and, above all, drone aircraft. Military research is investing huge sums in this field, and many people believe that war in the future will be ever more automated and fought at a distance, reducing the direct use of humans on the battlefield. As well as limiting the risk for humans, who would no longer be exposed to battle, military robots would have the benefits of endurance, unwavering attention and a lack of human feelings such as fear or compassion. From a technical perspective (which also has ethical implications), the much-feared risk is that the machines might become autonomous to a greater or lesser degree.

The use of robots for military purposes re-evokes all the problems already considered in relation to robotics in general, but also accentuates some crucial aspects. The inevitable prevalence of violent and destructive intentions over cooperative scenarios will lead to developments in the interaction between automated systems and human beings which will be unable to abide by the most traditional limits, not even the general prohibition of causing damage to objects and injury to people. However, the opposite position could also be argued: the conditions and precautions relative to the use of robots for civil objectives do not apply to military use, because if the aim is to kill, there is no difference between killing with a weapon or killing through a robot; between designing weapons to kill or arming a robot to kill. Once again, the problem is the underlying perspective. Are these machines like any other, or more accurately weapons like any other, whose aims and objectives fall within human modes of use? Or are we past the point of mere machines, and closer to armed entities whose degree of autonomy can be considered as approaching “transhuman” or artificial intelligence, hence opening up completely new legal and moral scenarios?

We must first distinguish between automated mechanical systems whose functions are limited to vigilance, control and assistance to combatants[23] and

[23] Drones can be used to locate and neutralise explosive materials (in Operation Iraqi Freedom in 2003, submarine drones were used for the first time to clear mines in the Port of Umm Qasr) and to scour and control the land (such as the Samsung SGR-1 sentinel robots used in South Korea, the Sentry Tech system used by the Israeli defence forces, and the Swarm vehicles used in the USA: these autonomous aircraft are capable of flying in synchrony like a swarm of

automated mechanical systems capable of killing (lethal weapon systems). Although the difference may be extremely subtle, as many defence systems can also be equipped to attack[24], it is undeniable that even if restricted exclusively to support of military operations, automation will be extremely useful in saving human lives and improving efficiency.

In their turn, lethal weapon systems are subdivided into human-in-the-loop (HITL) systems, which are remotely controlled by a human operator, and human-out-of-the-loop (HOOTL) systems (lethal autonomous weapon systems - "killer robots") which, once activated, are designed to interact with human beings and the environment without any human input. While HITL systems such as drones, which react to external stimuli under a varying degree of human control, are now widely employed in warfare, HOOTL systems, as far as is known, are still in the design phase. However, intensive research has already been conducted[25] into various degrees of automation which could enable discrimination between the different scenarios and hence development of the technology needed to move towards HOOTL systems in relation to limiting their collateral damage: frequency of contact with human operator necessary for correct function; ability to react and adapt to environmental uncertainties; safety and security in making decisions to complete the mission; ability to learn from unexpected events.

However, the idea of an intelligent autonomous weapon - a “killer robot” - could be misleading. Intelligent bombs are already “killer robots”, capable of following and destroying a target. However, they never decide to launch themselves, nor do they choose their target autonomously, replacing humans in the choice of enemy or war itinerary.

In warfare too, therefore, robotics and AI act as an extension or enhancement of what humans already do: in this specific case, an enhancement of fighting ability. Drone tanks and drone aircraft are already highly sophisticated examples of “killer robots”, without having to imagine some anthropomorphic being which would probably not actually be as dangerous as we fear.

From a technical perspective, it can be affirmed with some certainty that it is unlikely that an anthropomorphic robot could become a “killer robot”. The interconnections between the biomechanics and sensorial capabilities of the human body, and the human ability to improvise, are too complex to reproduce in a humanoid.

In any case, the rise in the use of drone weapons has led to numerous international conferences to discuss if the rules restricting the use of force set forth in the international law of armed conflict (LOAC) can still be applied to the various types of automated weapon systems. Over time, reflection on how to reduce abuse in military operations (*jus in bello*) has settled around a set of

---

bees and are designed to collect sensitive information and data). They could also have important medical uses, providing remote support and recovering the wounded. See J. C. Rossi, *La guerra che verrà: le armi autonome*, SIS, November 2016.

[24] All technologically advanced countries use or are preparing to use automated terrestrial or aerial systems to automatically identify and respond to missiles and rockets. Examples include the Israeli *Iron Dome* and the US’s *Phalanx Close-In Weapons System*. The British *Taranis* and US’s *X-47B* are aerial systems at an advanced planning stage.

[25] For example, by the National Aeronautics and Space Administration (NASA), the Air Force Research Lab (AFRL), and the United States Department of Defense (US DOD): R. TITIRIGA, *Autonomy of Military Robots: Assessing the Technical and Legal (“Jus In Bello”) Thresholds*, in “The John Marshall Journal of Information Technology & Privacy Law”, 2016, 32, p. 59.

principles that should guarantee, even in cases of the most extreme violence, respect of the principles of humanity (Martens clause). These principles of necessity, proportionality and distinction underlie international customs and are the basic premise of the various international declarations on the use of weapons, especially the four Geneva Conventions and their additional protocols. These principles should always be adhered to in all stages of a conflict: from the rules of engagement to modalities of combat, relying on the human capacity for weighing up the circumstances so as to limit the use of force to what is strictly necessary to reach the military objectives, clearly distinguishing between civilians and armed forces.

How far will it be possible to respect these conditions through the remote control of automated weapons whose efficiency is directly correlated with their ability to react immediately and autonomously? How could the LOAC, written with the assumption that it is humans who are fighting and killing other humans, restrict the action of killer robots? Until we can convert our ethical codes into machine codes, it will be impossible to dream up effective limits to the selective and destructive capacity of robotic systems, whose efficiency goes hand in hand with their lethality.

We must recognise that it is easier to program a destructive reaction than to build in an inhibitory mechanism. This cognitive difficulty throws a troubling light on our future. It is even more worrying that we have no idea if we will ever be able to overcome it. Mechanisms capable of distinguishing between different real-life situations are not yet within our grasp. From this point of view, automated weapons and armed robots give rise to a clear asymmetry which changes all the traditional parameters of the laws of war, because those arming the robots cannot be killed: the ethics of the battlefield shift towards the ethics of execution[26]. The operator’s personal risk is thus greatly reduced, transforming the "knightly" (albeit still atrocious) nature of a personal challenge into a kind of videogame where the confine between the imaginary and the real could become blurred, leading to further ethical unaccountability.

There is hence a risk of taking indifference towards death to extreme consequences. For a robot, following an enemy means recording an electronic signal, decoding an image and building a correlation between behaviour pattern and target (signature strikes). How can the distinction between civilian and military be programmed? Between children and adults? Between attack and surrender? Given how complicated it is to train robots to correctly interpret human gestures (body language), can we really let life or death depend on this interpretation?

The ethical problem of military robotics is essentially one of responsibility. Robots are a remote-controlled weapon. However, their control only goes so far: any weapon, however “surgically precise” or “intelligent” it might be, has a range of uncontrollable destruction, and the responsibility for its use falls on the entire operative chain, from design and manufacture, through whoever decided to employ it, down to the last operator to command or remotely control it. Up to this point, there is no difference from other more or less conventional lethal weapons. However, the alarming aspect of robotics is the potential margin of uncontrollability or autonomy which could arise through technological developments. We believe that all this must be foreseen and controlled, and it must in no way lead to a justification or a reduction of responsibility.

[26] G. CHAMAYOU, *Théorie du Drone*, Italian translation. Derive Approdi, Rome 2013.

Under optimal programming conditions, a robot may be "responsive", but certainly not "responsible". Responsibility for its use must always and exclusively fall on human beings, the only moral agents capable of answering questions on ethicalness. A mechanical or automated response, however complex and unpredictable, does not equate to a moral decision.

There is also a serious issue of transparency. One of the key elements in military responsibility for abuses committed in violation of the LOAC is reconstruction of the chain of command. Where robots are employed, the greater their autonomy, the more difficult it will become to identify a specific subject who can be imputed with a given action. Where does the responsibility of the programmer (who is asked to maximise the robot's efficiency in acting and reacting) end, and that of the person sending it into action begin? There is no sense in asking if drones violate human rights: they do not. It is not the drones which violate human rights, but the context in which they are used and which they themselves help determine.

All this makes it extremely urgent for the international community to take a position through an supplementary protocol to the Geneva Convention, in order to regulate and restrict the development of the dawning technology of autonomous weapons. Some nations (Ecuador, Cuba, Egypt, Pakistan, the Vatican State) have even requested their total ban[27].

The virtual aspect of the enemy, which is transformed into an algorithm which activates other algorithms, also has repercussions for police checks, negating the difference between military and civilian use of robots. The constant presence of threat necessitates the systematic and continuous surveillance of whatever images and communications are circulating the world. This surveillance is performed *a priori* on everything and on everyone, regardless of any specific situation of tension. If war is everywhere, so too must surveillance reach everywhere, with the suspension of some fundamental rights, and even making some rights such as the right to privacy or a fair trial unthinkable. While using robots for policing has undeniable benefits (greater reach, reduction of costs, opportunity to use human resources for other tasks), it could have a severe impact on the principles of freedom and privacy. This is even more the case if the data provided through robotic surveillance systems are crosschecked with other sensitive data and used to classify and profile the population. There is thus a need for legislation covering the conditions for the use of drones or any other automated system used for surveillance and policing purposes, requiring the judicial control of the acquisition, storage and use of images. In any case, the transparency of decisional processes and the public's right to information must be guaranteed. The use of lethal weapon systems for policing operations should be reserved for exceptional situations and authorised beforehand by the judicial authority.

Finally, both military and civil protection systems have problems in relation to security. Like all remote-controlled machines, they may malfunction or be "hijacked" by hackers or the enemy, thus failing to deliver the aims and interests of those who designed and programmed them.

In fact, it should be borne in mind that a connected society such as that of our future could be more seriously damaged by a hacker than a "killer robot". In theory, blocking the "internet of things" could halt the life of an entire country:

[27] There is even an NGO, *The Campaign to Stop Killer Robots*, whose aim is a total ban on lethal autonomous weapons systems (https://www.stopkillerrobots.org/).

traffic lights, railway points, dam spillways, air control systems, power plant safety systems, industries and so on could all be affected. Such a scenario could cause many more victims than an attack by intelligent weapons, and would require just one program.

Perhaps the most intelligent weapon of all is not found on the battlefield but is the one which "controls the control systems" of a digital society.

V. Legal responsibility

There is a need to identify, from a legal perspective, who is liable in the event that the function or malfunction of a robot causes injury to humans or, more generally, damage to the environment.

In the case of a mobile robot without the intelligence to be autonomous and capable of making choices, who has legal liability seems relatively simple: depending on the decisions of the individual legal system, it could be attributed to the manufacturer (designer, programmer, producer, team), to the vendor or to the owner or user. There could also be a shared or distributed liability. Compensation for damages, how it will be identified and calculated and what form of insurance should be obligatory can also be discussed.

Naturally, traceability mechanisms must be guaranteed, to enable the reconstruction of decisional processes and the machine's responsibility. This notion means that the manufacturer must ensure that the robot's decisional steps are subject to reconstruction, so that, from a technical perspective, the cause of all the robot's actions and omissions can be precisely determined by analysing its program algorithms. "A robot's decision paths must be re-constructible for the purposes of litigation and dispute resolution" [28]; in other words, this traceability is necessary to establish, from the robot's own data, what objective it was actually pursuing at the time it caused the damage or injury in question.

The traceability of the processes through which robots already make certain decisions (such as in finance) also has a fundamental role in identifying any faults in the reasoning behind the robot's decisions, which must be taken into account when establishing compensation for any violated rights.

The European Parliament's Committee on Legal Affairs, in its previously cited project, also underlines the importance of traceability: "For the purposes of traceability and in order to facilitate the implementation of further recommendations, a system of registration of advanced robots should be introduced, based on the criteria established for the classification of robots"[29].

However, as already noted, we are still a long way from being able to create robots capable of making decisions and acting upon them independently of external control. Robots cannot yet be considered as "electronic people" in the sense of moral agents. In fact, for the reasons already discussed they are not even "moral patients", because only living beings are capable of suffering, not robots, which have no emotive capacity; even if they can simulate emotions, they can feel neither pain nor pleasure. If they are not moral agents, they cannot be considered as such from a legal perspective, and therefore the

[28] L.D. RIEK, D. HOWARD, *A Code of Ethics for the Human-robot Interaction Profession*, in *Proceedings of We robot*, 2014, p. 6.

[29] Report with recommendations issued by the European Parliament Committee, cit., p. 14.

anthropomorphic traps which limit discussion to the traditional ethical and legal categories must be avoided, to enable new categories to be invented.

Today, therefore, we are far from having robots whose nature enables the attribution of direct liability. However, as it cannot be definitively excluded that enhanced autonomous decisional abilities may be a feature of robots in the future, it is advisable to considered this possibility now.

We therefore believe that the possibility of having different judicial categories to those already existing for other mechanical devices (cars, aeroplanes, etc.) must necessarily take account of the amount of control transferred from the programmer or owner to the robot itself and whether the latter has sufficient AI to enable it to make its own decisions and react to its environment, thus revising its choices and positions in relation to its programming. However, it is already possible to envisage new risk and liability categories in relation to the unpredictability and growing autonomy and learning capacity of robots. An obligatory insurance scheme seems necessary, as does a fund to ensure that damages can be compensated for, to which all parties (manufacturer, programmers, owners, users) could contribute[30].

Unlike other sectors, in which technological development produces “objects” used by humans, in this case “subjects” with AI and largely autonomous of humans will be created. Given concerns about the possible impact of robotics on human safety and security, private life, integrity, dignity and autonomy, it can be seen that applying to robots the judicial schemes and doctrines created for the traditional legal categories (natural persons, legal entities, chattels) would be inadequate. Moreover, numerous social and operational sectors (especially the medical, welfare and military sectors) necessarily affirm that the advent of robots necessitates a code of conduct which is both as universally recognised as possible and legally binding, in the interests of both individual citizens and businesses.

VI. Recommendations

The Bioethics Committee and Biosafety Committee, aware of the complexity of this continuously evolving technology, hereby offer some recommendations on robotic applications in social, medical, military and legal contexts.

<u>In social contexts:</u>

1. The importance of critical information for the public on the developments, potential and limits of robotics and AI to enable a critical awareness and avoid excessive enthusiasm or rejection influenced by unrealistic scenarios worthy of science fiction, and the importance of studies to help train citizens in the development of the essential technological skills for the era of robotic revolution.

2. The desirability of an interdisciplinary analysis of the impact of robotics on society (and on labour in particular), with attention to its psychological and emotive impact, and the implementation of strategies to avoid dependence on robots and the replacement of humans with machines, with the aim of maximising the value of human labour in the robotic age.

[30] See also the recommendations of the Report of the European Parliament Committee, cit., p. 12.

3. The key consideration of the robotic divide and the importance of both avoiding discrimination between the included (those already living in a technological and robotic society) and the excluded (through an inability to gain the necessary skills) and promoting ways to assist people in conditions of "technological vulnerability" (the elderly, people with cognitive disabilities).

4. The need to implement ethical codes for robot programmers and the need for ethics committees for robotic research which facilitate interdisciplinary discussion between scientific, ethical and legal experts in the context of rapid robotic innovations.

5. The importance of introducing the study of ethics into engineering and IT courses, to encourage moral reasoning skills in the context of new robotic technologies from the very start of training.

<u>In medical contexts:</u>

1. To promote adequate experimentation of robotics in surgical and welfare settings so as to guarantee conditions for the user's physical and psychological integrity, explaining the risks and benefits, including in the informed consent to the use of robots.

2. To ensure equal access to robotic technologies and that the use of robots is to assist and not replace humans, so as to avoid delegating to machines the irreplaceable human duty of care and assistance.

3. The necessity that the introduction of robotics in medicine should always involve consideration of the real benefits, the complexity of the total change to the service facility and the financial burden this carries.

<u>In military, policing and surveillance contexts:</u>

1. The need to conduct studies on ethical issues in military robotics, highlighting the limitations and consequences with regard to the principle of responsibility in relation to humans/autonomous machines.

2. The urgency for the international community to take a position through an supplementary protocol to the Geneva Convention, in order to regulate and restrict the development of the technology, currently in its infancy, of autonomous weapons.

<u>In legal contexts:</u>

1. In relation to the legal liability of robots, to put in place, with immediate effect, safeguards and guarantees for the public, for users and for businesses, to avoid as far as possible that the conduct of robots might cause injury or damage, considering that the aim is not to create laws for intelligent robots but to create rules for the humans relying on these machines to enhance their own abilities, whatever they might be.

2. To establish, in any case, cover for the injuries and damage that robots might cause to users, the public or the environment, also taking account of the robot's autonomy and learning capacity and thus the control that has been transferred from the programmer or owner.

3. A unifying legislative standard, at least in the European Union, is desirable to ensure legal consistency and certainty.