18

Technological Changes and Rights Evolution in the Bio-digital Era

A Philosophical Overview

Marta Bertolaso and Alfredo Marcos

18.1 INTRODUCTION

Technology, converted into technoscience in the course of the last century, gives us more power than simple traditional techniques and therefore burdens us with greater moral responsibility. This responsibility becomes even greater when we can technically modify the foundations of life with new *biotechnologies*. That is, our moral responsibility grows progressively as we technically modify the human being – the moral subject – either by technologically changing human biology or by conditioning human behaviour and the place of humans in the world and society. There are currently two major *anthropotechnical proposals* to confront with the challenges posed by technoscience and current biotechnologies: the *bio*-project and the *info*-project. The first tries to modify the human being through biotechnologies; the second aims to modify, complement, imitate or even replace human decisions and behaviour through info-technologies. In light of these projects, we can say that we have entered the *bio-digital age*, which raises new *ethical challenges* and asks for a deeper reflection on the foundations of human rights and the possible scenarios of their evolution.

In this chapter, we present such challenges and offer a critical reflection on them, as well as some proposals for considering the bio-digital era from a humanistic perspective. In Section 18.2, therefore, we outline the main roots of the emerging ethical challenges. In Section 18.3, we address the bio-challenges relying in part on examples in the other chapters in this volume. In Section 18.4, we focus on the philosophical assumptions and implications of the major info-challenges. Here we identify and discuss ethical questions that emerge at the crossroads of the fields of human-technological development related to the degree of autonomy and of dependencies mediated by machines. In Section 18.5, we conclude that improving human life to make it more properly human requires a new consideration of human solidarity and care open to new scenarios. Such new scenarios would take into account the first generation of human rights and their potential to evolve within a framework of a *technological humanism*. By 'technological humanism' we mean a specific way of thinking about the technical that would require technological developments to be at the service of improving human life. We should not be aiming with our technologies to create superhumans, transhumans or posthuman beings. We should not be seeking to modify our common human nature, but rather to support the flourishing of human life. New technologies can help us live as human beings. making our lives more properly human.

18.2 BEYOND TECHNOLOGICAL ADVANCES: THE ROOTS OF THE EMERGING ETHICAL CHALLENGES

In this section we review the new anthropotechnical proposals from the perspective of the technological humanism we support. Human beings are technical beings. Their lives cannot be accomplished without the technical. We depend on it to eat, to communicate, to learn. Our bodies have been co-evolving with our techniques. The structure of our chewing apparatus, for example, is correlated with our mastery of fire and our capabilities for making and using stone tools. A rejection of the technical would therefore be a rejection of the human. A radically anti-technical attitude is in discord with human nature and human life, since the human being is by nature a technical being. At the same time, an acritical acceptance of technological development as such seems inadequate for the human future. Legal and ethical questions clearly arise at the crossroads of these two positions. Thus, while it is correct to ask under which conditions the technical might become a threat to human life and fundamental rights, it is also important to ask under which conditions the technical favours human life and facilitates the full realization and flourishing of human beings.

This 'technical' occurs in a number of modalities (traditional techniques, technology, technoscience, biotechnology, anthropotechnology). To appreciate the specificity of each of these modalities, it is useful to observe how the technical has developed throughout human history, starting from *traditional techniques* (as in, for instance, the manufacture of flint axes or pottery), and how more recent technological applications and scenarios have appeared. The technical, when it joined science at the beginning of modernity gave rise to *technology* (steam engines, electronics, television etc.). Francis Bacon's work is, perhaps, the classic *locus* for understanding the foundations of this confluence of technique and science. We find here a second modality of the technical, which differs in many aspects from traditional techniques, giving us more power, better opportunities and, at the same time, exposing us to greater risks. By the end of modernity, the symbiosis between science and technique had become so tight that we began to speak of *technoscience* (nuclear physics and engineering, computer sciences, robotics etc.). In this chapter, we focus specifically on this third modality of the technical. Technoscience, by further increasing powers and risks, burdens us with a greater moral responsibility. The interesting assumption beyond it and resulting from the co-evolution of the empirical sciences' knowledge and technology is that complete control over nature, including ourselves, will become more and more possible. This scenario dominates current debates and concerns from science to economics and finance.

Returning to the path of evolving modalities of the technical, we move on now to the fourth type, *biotechnology*. Our moral responsibility becomes even greater when we begin to modify living beings by biotechnological means (gene editing, for example) which affect natural processes in a very profound way. Our moral responsibility grows even more when the living beings we are trying to modify are human beings, that is, moral subjects, those responsible for technological agency.

The fifth and most radical modality of the technical is *anthropotechnics*, applied directly onto the human being. An atomic artefact can be powerful and destructive. In a sense, it can be considered much more powerful than a modest genetic modification. But if that genetic change affects human beings, then we are modifying the creators and potential users of the atomic artefact itself. Genetic engineering can modify, for better or for worse, the human being – the very being who makes, suffers and uses (or not) atomic bombs. Contemporary anthropotechnics

multiply by a previously unknown factor the opportunities, risks and responsibilities derived from the different modalities of the technical.

Summarizing, we can currently acknowledge two major anthropotechnical proposals. These are the *bio* project and the *info* project. The first tries to modify human beings through biotechnologies, and the second aims to modify, complement, imitate or even replace human beings by means of info-technologies. We can therefore say that we have entered the *bio-digital age*, which raises new *ethical challenges*. In Sections 18.3 and 18.4 we discuss examples of such proposals in more detail. In Section 18.5, which concludes this chapter, we present related challenges through a critical reflection, along with some suggestions for tackling the bio-digital era without losing a humanist perspective.

A further clarification is needed before delving into the ethical questions raised by the new anthropotechnics. The dual disposition of anthropotechnics (bio and info) suggests a dualistic anthropology that has been reinforced by different forms of scientism, so that we still suffer a reduction of the human questions about truth to statements of factual and empirical evidence mediated by measurements whose limits are shown by control paradigms in the current scenarios. This means that the current main lines of anthropotechnical intervention, the *bio* and the *info*, are not accidental but rather respond to the aforementioned anthropological conception, which follows the techno-scientific transitions that we will now briefly and critically examine.

The most remote precedents of this anthropology are found in Platonic and Gnostic thought. Its modern precedents reside in the Cartesian philosophy of the two substances, *res extensa* and *res cogitans*, as well as in the two basic realities that Kant identifies: nature and freedom (the famous starry heavens moral law of his *Critique of Practical Reason*). Ancient and medieval philosophy, in their Platonic aspect, have tended towards dualism, although not in its Aristotelian aspect. Modern philosophy, also, has been marked by a dualistic anthropology. It is not surprising that anthropotechnical projects respond to this way of thinking. Some seek to change human beings by profoundly modifying their bodies, at either a genetic or neural level. Others try to separate the human spirit, now conceived as a kind of software, from the vicissitudes of the body, to free it from the weakness and finitude of the flesh, digitize it and upload it to a cloud, or completely replace humans with some other genre of disembodied digital entities, such as an artificial intelligence system.

Ethical problems of the new anthropotechnics thus also derive from a deeper, anthropological and metaphysical problem. Consequently, to ethically judge the different anthropotechnical proposals and scenarios we are faced with we need to elaborate a much more sophisticated understanding of human beings and of our specific way of being, belonging and dwelling in the human natural environment that, given the aforementioned premises, also includes the artificial. From this perspective, the ethical questions and challenges arising from the increasing discussion about human rights can be better and more fruitfully framed, for science and technology themselves can be better understood in their potentialities at the service of human beings' ontological unit, which is relational and societal (we learn with and through others, we realize ourselves by joining and dwelling in the real world as a community, and we also need others for our own happiness). In the following sections, we will examine how this happens in practice by analysing the ethical challenges that bio-anthropotechnics and info-anthropotechnics present. As stated in this section, in both cases we will consider questions of a metaphysical nature, without which the ethical approach would always be precarious.

18.3 BIO-CHALLENGES

We have already shown the progression that takes us from traditional techniques to anthropotechnics, through technology, technoscience and biotechnology. We have also seen the two main avenues of current anthropotechnical projects, *bio* and *info*. We will now explore the ethical implications of bio-anthropotechnics. The key term for discussing bio-anthropotechnics is *human enhancement*,[1] the intention of which is to improve the human through the application of various biotechnologies. Needless to say, there are approaches to human enhancement that do not involve biotechnology (such as culture, therapy, training etc.), as well as biotechnologies that do not contemplate contributing to human enhancement. At the same time, however, there exist or have been projected a number of biotechnologies whose main goal is human enhancement. Although these technologies can affect any part of our bodies, for example in the form of prostheses or implants, most of the proposals for technological intervention are directed towards the human organism's two main sensitive systems: the (epi)genetic system and the nervous system (see also in this volume Ligthart, Meynen and Douglas (Chapter 3), Andorno (Chapter 19), Bennett and Richards (Chapter 9)). Interventions aimed at the most profound improvements in our bodies involve implanting or using devices affecting these two systems.

The illusion of improvement has been fuelled in recent times by the public communication policies of mega-research projects. We consider this an illusion because if human nature were modified, the evaluation criteria would be lost. For this reason, one cannot properly speak of improvement. It should be remembered that *big science* – that is, large-scale projects usually funded by groups of national governments and big companies – was born with the Manhattan Project in the middle of the last century, but its full measure was perceived, in the final stretch of the century, by the Human Genome Project (HGP), the conclusions of which were published in 2003. Then it was the turn of projects on the human brain (BRAIN and Human Brain Project). Such large scientific projects depend critically on media, political and financial factors, which tend to incline their promoters to formulate (or at least suggest) dazzling promises and expectations, such as human enhancement or state of health predictions and risk calculations (see also in this volume Gunnarsdóttir, Cohen, Minssen and Gerke (Chapter 4), and Sandor (Chapter 5)), often based on the aforementioned inherited philosophical assumption about science, anthropology and technology.

Therefore, the agendas beyond the applications of such advances often show two very different interpretations of the empirical data and its relevance in terms of knowledge, and thus its normative relevance from a legal or ethical viewpoint. On the one hand, there are those who think human beings are completely natural, that is, determined by empirical laws and dynamics. As such, they can be improved in terms of their performance and functions in many ways. The interventions that are projected from this naturalistic anthropology try to turn human beings into 'better' living beings, as we do with plants and animals whose 'good' is merely a result of our cultural, commercial and pragmatic aims. Although they try to take charge of evolution to obtain a new and better type of organism from the individuals of our species, rather than improving human lives, some agendas seek to transform and overcome the human, that is, to build a post-human being that would be invulnerable to the limitations and ailments of our species. Within this line of thinking, transhumanism (TH) is currently the most influential school of thought.

[1] M. Bertolaso, 'Le Human Enhancement Technologies e l'Irriducibilità della Complessità Biologica' in: S. Kampowski, D. Moltisanti (eds.) *Migliorare l'uomo? La sfida etica dell'enhancement* (Siena: Cantagalli, 2011), pp. 35–58.

On the other hand, biotechnologies can be applied also in projects for the *improvement of human life*. From the viewpoint that each human being, each person, already has an infinite value that we call *dignity*,[2] it is not possible to improve, overcome or add value to any specific person through biotechnological interventions. The aim then becomes to improve human life, to make people's lives better. This means making people's lives more *properly human*. In other words, human nature has a normative function. Thus, a properly human life will be a life according to human nature. If we understand that human beings are social and rational animals, then a properly human life will be a life in which these three aspects (social, rational and animal) can be developed in a harmonious and integrated way, a life with physical and mental health, in peaceful and just coexistence, with possibilities of learning and spiritual growth, with a certain freedom, personal identity and capability of recognizing the same characteristics of freedom and identity in others. Biotechnology applications can contribute to such improvements of human life, and do so, for example, when they have a clearly therapeutic purpose (e.g., see in this volume Pierce and Villaronga (Chapter 6), Brotugno (Chapter 8)).

Currently, and regardless of the intrinsic value of these two different anthropological projects, the first one is moved by an intellectual initiative, thanks to the contribution of the transhumanist, while the second one remains in a defensive and critical position. TH has become, in fact, an authentic cultural trending topic and has set the dominant philosophical agenda. There are those who argue, for example, that future policy will be structured according to attitudes towards anthropotechnics, and TH has already taken a position in the debate leading to the most favourable proposals for anthropotechnics.[3]

Among transhumanist thinkers, the Oxford nucleus stands out, and within this Julian Savulescu and Nick Bostrom are prominent.[4] In the United States, Max More founded the *Extropiano Institute* in 1990, devoted to promoting TH. We owe to More[5] this definition, 'Transhumanism is a class of philosophies that seek to guide us into a post-human condition.' We recognize here a clear desire to overcome human nature, understood in terms of constraints, never of constitutive limits. There is, in addition, another front in defence of deep anthropotechnics made up of continental post-humanist philosophers, among whom we can name, for example, Peter Sloterdijk and Giorgio Agamben, whose inspirations must be sought in the existentialist and nihilistic tradition.

The aspects in which the supposed enhancement of human beings is thought possible range from physical, moral and intellectual capacities to emotions, longevity and the prolongation of youth.[6] As discussed by Andorno (Chapter 19) in this volume, the trend is acquiring further importance given the possibilities of human germline engineering, posing important questions regarding how in the future equality, as the quintessential notion of human rights, can be assured.

TH, in its different forms, has been criticized on many fronts. Some people think we are facing mere futuristic daydreams without viability. There are also doubts about the social viability and sustainability of a greater human longevity than is currently known. What if each person lived three or four hundred years? Some argue that, under the precautionary principle,

[2] R. Spaemann, *Love and the Dignity of Human Life* (Grand Rapids, MI: Eerdmans, 2012).

[3] A. Marcos, '*Bases filosóficas para una crítica al transhumanismo*'(2018) 7 *Artefactos. Revista de estudios de la ciencia y la tecnología* 107–125.

[4] J. Savulescu and N. Bostrom, *Human Enhancement* (Oxford: Oxford University Press, 2009).

[5] M. More, Transhumanist (2010). FAQ, www.extropy.org/resources.htm/

[6] See also A. De Grey and M. Rae, *Ending Aging: The Rejuvenation Breakthroughs that Could Reverse Human Aging in Our Lifetime* (New York: St. Martin's Press, 2017).

there should at least be a moratorium on some anthropotechnical projects. On the other hand, following social thought, one can doubt the justice and equality with which the alleged enhancements would be made. Politically, both Jürgen Habermas[7] and Francis Fukuyama[8] have already insisted on the dangers of the transhumanist mentality for freedom and democracy. Also in the field of bioethics, there has been criticism of a project of alleged human enhancement. The principle of autonomy has been questioned.[9] As for the principles of beneficence and non-maleficence, it is more than doubtful that uncertain experiments on humans can be accepted without any clinical gain. The principle of justice would also be compromised to the extent that the resources available for an enhancement compete with those dedicated to strictly therapeutic actions.

Our position is aligned with that of those who take a critical attitude towards applications of the new technologies on humans and human environments. It is, in fact, possible to challenge TH from the perspective of the very concept of human nature[10]: a practice that tries to go beyond human nature destroys in passing the evaluation criteria. Consequently, it can hardly be called improvement or enhancement. It would be, in principle, a change without an axiological component. And basically, it would be a change for the worse since the loss of the axiological component – indifference – already supposes a loss of value. To paraphrase Hans Jonas,[11] we could now say that the transition to the impossibility of value is already a disvalue.

We endorse going back to the fundaments of human dignity and the main features of our natural life and scientific knowledge that characterize human beings as dependent, rational and social beings. We understand human dignity as Kant defined it, as a special and absolute value based solely on our humanity. It has nothing to do with class, race, gender, abilities or any other factor other than being human.[12] There are also objections to TH agendas and strong anthropotechnical cultures from the epistemological and ecological perspective, since the very nature of complex sciences (such as biomedicine, ecology, climate sciences or epidemiology) does not allow for possible effects and risks to be predicted in a deterministic and mechanistic way.[13] The possibilities for living systems to find different ways to reach their goals of survival, reproduction and repair at the different levels of the biological organismic or eco-systemic organization, are commonplace. Uncertainty and multiple realizability limit the possibility of controlling living systems as we do with machines. Current examples can be easily found in the debates about the COVID-19 pandemic and the real possibility of finding a unique and definitive solution in a vaccine.

As is easy to understand, this implies the assumption that any real challenge is a human, anthropological one, that is, which is scientific, technological and ethical at the same time. This in turn implies overcoming the dualisms and reductionisms mentioned previously as philosophical foundations of major anthropotechnic agendas. When, instead, the emphasis is on the risks or badness of technology per se, we argue that the question and the overall debate is simply badly

[7] J. Habermas, *The Future of Human Nature* (Cambridge: Polity Press, 2003).
[8] F. Fukuyama, *Our Post-Human Future: Consequences of the Biotechnology Revolution* (London: Picador, 2003).
[9] Habermas (note 7).
[10] L. Kass, *Life, Liberty and the Defense of Dignity: The Challenge for Bioethics* (San Francisco: Encounter Books, 2002); M. Sandel, *The Case against Perfection: Ethics in the Age of Genetic Engineering* (Cambridge, MA: Harvard University Press, 2007).
[11] H. Jonas, *The Imperative of Responsibility* (Chicago: University of Chicago Press, 1985).
[12] A. Marcos, '¿Tienen igual dignidad todos los seres humanos?' in: M. Pérez de Laborda, C. Vanney and F. J. Soler (eds.) *¿Quiénes somos? Cuestiones en torno al ser humano* (Pamplona: Eunsa, 2018), pp. 210–214.
[13] M. Bertolaso, N. Di Stefano, G. Ghilardi and A. Marcos, 'Bio-Techno-Logos and Scientific Practice' in: M. Bertolaso (ed.) *The Future of Scientific Practice: 'Bio-Techno-Logos'* (London: Pickering & Chatto Publishers, 2015), pp. 179–191.

framed and posed. In light of this, as we will further develop in Section 18.5, care and the possibility we always have of caring for others emerge as an interesting and actionable viewpoint that should be more and more endorsed for politics and normative decisions.

18.4 INFO-CHALLENGES

In this section we focus on how the deployment of digital systems, both for the automation of processes and for the implementation of information and biological performance, is posing new and challenging questions for our understanding of human beings, their relationships and social dynamics.[14] Specifically, the use of machine learning/deep learning algorithms associated with Big Data mining/processing in health and educational processes poses important challenges to guaranteeing values such as equality, freedom of behaviour, protection of privacy and intimacy (see in this volume Rotenberg (Chapter 12), Bygrave (Chapter 13), Valcke and Verschaeve (Chapter 14)). The relationship of info-technologies with human life is, in fact, quite complex. Sometimes they are simply presented as tools to improve human life. This is the case posed by Botrugno (Chapter 8) in this volume about the spread of telemedicine services in contemporary healthcare systems that is reshaping the medical practices and organizational patterns of healthcare delivery.

However, on other occasions more daring claims are made about info-technologies, such as their ability to simulate human capabilities and functions, overcome human beings or even become a substitute for them. A crucial example which spans bio and info technologies is presented by Douglas on the application of persuasive technologies for 'smart rehabilitation'.

Such examples become more interesting as they progress in projecting expectations mediated by the application of 'artificial intelligence' (AI). Here the term refers to the supposed intelligence of some artificial artifacts or systems.

As is well known, the term 'artificial intelligence' was coined in 1956, during the Dartmouth Workshop (*Dartmouth Summer Research Project on Artificial Intelligence*), by the American scientist John McCarthy. It is true that, from a rhetorical and propagandistic point of view, this term has been a complete success. But the term 'intelligence' refers to the ability to understand. We know that an artificial system is incapable of understanding. It cannot even properly be said that a machine counts or computes. Counting involves bringing together two (or more) different moments, and keeping them as such, in a single conscious representation, which a machine does not do. On the other hand, it is true that AI can help us solve problems (computation, geolocation, logistics, telephone assistance, assistance with medical diagnosis, advertising and a long etcetera). Sometimes intelligence is even defined as the ability to solve problems. However, these are not problems for the artificial system but rather for the designer or its user.

For a facial recognition system, whether or not to recognize offenders is not a problem. Recognizing offenders is a problem in relation to people's safety, and the system can help us deal with it (see in this volume Garvie (Chapter 11)). Of course, the same system can serve to follow the movements of the population of a country and to facilitate political repression. But this is not a problem for the cameras or for the software involved. It is, undoubtedly, a problem for the inhabitants of such a country, victims of political repression. None among hammers, abacuses and the most advanced computer systems have problems. The problems as such are ours, as living beings and as human beings. Only a living being can die or suffer, and only people can wonder about the meaning of their lives. Causes of death and suffering are problems. Both a

[14] A. Marcos, 'Información e Inteligencia Artificial' (2020) 12 *Apeiron: Estudios de Filosofía* 73–82.

hammer and a computer network, each in its own way, can help us deal with them (or make them worse) but this does not make them smart.

There are, of course, other alternative characterizations of intelligence, for example as information processing or as an ability to perceive and adapt to the environment. However, in one way or another, all of them refer to the human being, given that information is a triadic relationship which implies an intentional consciousness (*a* informs *b* on *c*). On the other hand, only living beings *perceive*. When we say that a camera or a motion sensor perceives, we are obviously using this term in a metaphorical sense.

This distinction is clear also from a different perspective. AI is sometimes characterized by its simulation capabilities. It is said to simulate functions of human intelligence, but simulating is not *being*. Simulating intelligence is not the same as being intelligent. Furthermore, the notion of function inexorably refers to a being for which or whom a given effect is functional, although artificial systems also depend on the functionality they may have for humans. Outside the human context, the lights that go on and off on a screen or the movements of a robot are mere effects. They do not fulfil functions.

Even more so, the concept of *data*, like that of a *message*, implies the concept of a receiver of the data, and the receiver, directly or indirectly, will be a conscious human being. It is our gaze on our computers which transforms into data for our computers all the inputs that come to them, and into data for us all the outputs they emit. Only by assuming a consciousness in the background can we speak of certain states of matter as data, and of an AI system as a receiver of data. As Charles S. Peirce wrote, 'By semiosis I mean [. . .] an action or influence which is or involves a cooperation of three subjects, such as a sign, its object and its interpretant, this three-relative influence not being in any way resolvable into actions between pairs.'[15] The concepts of *data* and *message* are both notions that belong to the field of 'semiosis'. They imply, as Peirce points out, intentional triadic relations. Thus, an electromagnetic change occurring in a computer is not by itself a *datum*, unless a conscious being can refer that change to a state of affairs external to the electromagnetic configuration itself. Let us think of the state of a computer when it records data about the sale of a certain product. This state (sign) is about the sale of the product (object) if and only if an 'interpretant' intentionally connects both events, being otherwise a simple physical state with no meaning at all.

There is thus an ontological difference between humans and machines that justifies framing the hope (or the threat) of cancelling such difference through technological sophistication as illusory, a mere categorical mistake. Artifacts, in the Aristotelian tradition, are substances only in an accidental sense, by analogy. Living beings, and especially human beings, that is, people, are substances in their own paradigmatic sense: the bottom-line question is an ontological one.

Strictly speaking, as Luc Julia,[16] one of the creators of the AI system SIRI, states, 'artificial intelligence does not exist'. How has the claim of the existence of intelligent artificial systems come into being and why does it have such an impact in our discussion about ethics and human rights? With the rise of personal computers, the miniaturization and lower cost of components, and the arrival of the Internet, both computing capacity and the contribution of data skyrocketed in a few years.[17] Some companies soon saw in this take-off a promise of abundance, and their

[15] C. S. Peirce, *Collected Papers* (Cambridge, MA: Harvard University Press, 1932), p. 484.
[16] L. Julia, *L'intelligence artificielle n'existe pas* (Paris: Éditions First, 2019).
[17] M. David and C. Sauviat, *Intelligence Artificielle* (Monaco: Éditions du Rocher, 2019), pp. 69–71.

own activity generated even more data and greater capacity to process it. Google, Amazon, Facebook, Apple and Microsoft (collectively known as GAFAM), as well as the Chinese Baidu, Alibaba, Tencent, Xiaomi (BATX) and Huawai, have rekindled the expectancies for AI, and they have done so thanks to *big data* and an exponentially growing and increasingly affordable computing capacity.

Thanks to this rebound in AI research, very useful and precise systems have been created in various fields. These innovations are always related to systems specialized in a certain function. Undoubtedly, machine translation systems; Internet of Things solutions and control systems; facial and visual recognition systems; expert systems in biomedicine and finance; automatic driving systems; and robots that simulate conversational skills (*chatbots*) have improved. None of these systems understands what they do, but that does not diminish their functionality . . . for us. In spite of all the achievements in the field of AI, no strong or general AI system – that is, an artificial system that can cope with all the functions that human intelligence performs in an integrated way – exists. The projection, however, of scenarios in which they do exist is a real expectation of further technological development. These projections, however, avoid the question of what human capabilities should be equally developed to maintain our mastery over technology, to preserve our behaving in a human way, that is, rationally pursuing our goals and aims and personally/collectively caring for the natural world and for others.

A similar reflection arises when we think of examples in which AI enters into more specifically decision making processes. In the field of clinical medicine, AI helps physicians decide whether or not they should consider a specific diagnosis. Similarly, in the field of finance, it helps bankers decide whether or not they should grant a certain loan. The question is who or what really makes the decision? Will the AI system make decisions for us? Is the human financial adviser superfluous? Is the doctor's opinion superfluous when an AI system diagnoses cancer? Is the human pilot superfluous when an intelligent navigation system decides to vary the speed or altitude of flight? Can acting against the output of an AI system always be reckless? What or who is responsible in each case for a possible failure? Should we mitigate with quotation marks the meaning of 'decide' and 'diagnose' when they refer to an AI system function? Similar questions emerge in relation to anonymization and transparency processes (see Garvie (Chapter 11), Rotenberg (Chapter 12) in the volume). As Rotenberg's analysis clearly shows, algorithmic transparency is now firmly established as a fundamental right and the cornerstone for the regulation of Artificial Intelligence. But he also rightly discusses how

> [l]egal instruments are not self-executing. To establish and maintain the right of algorithmic transparency, individuals and organizations will need to pursue legal claims [. . .] It is my view that algorithmic transparency is the most urgent goal for AI regulation and that it will become a critical indicator of the health of democratic societies in the years ahead.

This affirmation raises new questions. Given that the requirements of auditability and transparency can run up against the insurmountable problem of time and that AI systems are used to save time, can we, the humans, find the time to audit tasks that such systems perform at breakneck speed, or will we end up entrusting other AI systems with the audit and supervision tasks and so on?

Similarly, autonomy and data privacy, data quality and intellectual property become the fundamental criteria from which ethical and legal implications arise when AI systems provide recommendations for researchers who wish to integrate citizen scientists and Internet of Things devices into their research (see Scheibner, Jobin and Vayena (Chapter 16) in this volume).

We immediately glimpse the philosophical, anthropological, social, legal, economic, moral, labour and other implications.[18] The list of concerns and risks, in fact, is doomed to go well beyond any manageable measure or scenario if we discard these concerns and neglect to develop an adequate philosophical and anthropological framework to deal with these new challenges, through a proper understanding of information~~, understanding~~ and decision processes.[19]

AI works well as a complement to human intelligence. It takes on entity and meaning only in the framework of human action. The first and most obvious reason is that it originated as a product of human action. But there are those who think that once started it could self-maintain, even self-improve beyond the human limits. However, as said, the problem regarding the limits of AI is not technical but ontological. The truth is that there would not even be data (or big data) in the absence of a human receptor, of a person who is capable of unifying in consciousness the dispersion of the gross facts, of giving intentionality to each fact, thus connecting some parts of reality, which figure as data, with others about which these data tell us something. As we have explained, based on Peirce's ideas, data are semantic entities, that is, they are triadic relations (sign-object-interpretant). There is a *datum* when there is a consciousness that establishes an intentional relationship between two events: one that is taken as a sign and another taken as its object.

In addition, however lax we may be in the selection of data, there will always be some selection, since we cannot feed a machine with *all* the data that exists. In collecting them and in their contribution to AI systems, we continually make inexorably human judgements of relevance.

From these ontological considerations, we can derive an idea that is crucial for ethical issues: only people can make decisions, and such decisions should take into account more and more the relational and rational (not merely intelligent) nature of human beings. This idea derives from the basic definition of 'decision': 'a choice or judgement that *you* make after thinking and talking about what is the best thing to do' (Oxford English Dictionary, our italics) The mental or personal aspect of the concept is clearly acknowledged in its definition. If a machine has no mind (as is the case), by definition it cannot make decisions. If we remove the mental condition from the concept of 'decision', then we are talking about something else, not about what we normally understand by this term.

The very concept of decision is foreign to the mechanical. What we call 'decision' in an AI system will be so only to the extent that a human being has made the genuine decision to delegate some action to the system, that is, to automate it. The ultimate responsibility, whether things turn out right or wrong, can only be a human being's, and such responsibility rightly emerges from the interest and concerns of humans for each other's (collective and social) rights. It follows then that it is our responsibility to rely on the best AI systems available when making decisions. We also need to develop a stronger awareness about what our social and relational nature asks for. In many cases, it will be convenient for a doctor to rely on an AI system to support a diagnosis, especially if the system is well calibrated and has been shown to be useful and reliable in clinical trials. But, in all cases, the human expert must be free to act even against the indications of an AI system. In many cases we will need *someone* who understands, who creatively searches for new explanations, who reconsiders and generates new systems of expectations.

[18] Ibid.

[19] M. Bertolaso and F. Sterpetti (eds.), *A Critical Reflection on Automated Science: Will Science Remain Human?* (Dordrecht: Springer, 2020).

18.5 HUMAN RIGHTS AND TECHNOLOGICAL HUMANISM

Throughout the previous sections we have suggested that a philosophical reflection on the importance and relevance of technical power and responsibility is crucial to safeguard the duties of future generations (following Hans Jonas), that is, to safeguard for them a life which is properly human. Today, more than ever, we need a philosophical reflection on the impossible detachment from technology of human beings (following Heidegger), a philosophical reflection that overcomes a dualistic anthropology. On the other hand, an ontology of the digital cannot exist beyond a crudely physical one if we omit the human factor and the typical viewpoint of human beings on technology. The current bio-digital transitions, while shaping structural and institutional aspects of human life, are forcing us to deepen the foundations of human rights to better account for what it actually means to be human today, embedded as we are in a novel socio-technological trend.

Improving human life to make it more properly human requires a new consideration of the care and caring dimensions of human work,[20] of dependence-autonomy relationships. It also requires the creation of new scenarios that consider deepening the potentiality of first-generation human rights or that move towards fourth-generation human rights for AI and biotechnologies.

Since the 1980s, the distinction of three generations of human rights has become common. It reflects the chronological order in which they have been made explicit. Those of the first generation are most fundamental, since they include the right to life and liberty. Second-generation rights have an economic, social and cultural character. They include, for example, the right to a job or decent housing. Third-generation rights go further and include, for instance, the right to a healthy environment. Today there are advocates calling for a fourth generation of human rights. Fourth-generation rights would be especially related to the development of new technologies. They would include, for example, so-called digital rights, such as the right to a digital identity. From our point of view, it is not necessary to invent a fourth generation of rights to correctly manage the impact of new technologies. We believe that it is more promising in this regard to fully develop and apply the most fundamental human rights.

We thus have aimed here to offer a critical map of the aforementioned issues that can serve for the development of practical tools and criteria for decisional processes at a personal and institutional level. Moving from epistemic considerations on the limits posed by the intrinsic nature of machines which are limited by their inductive nature, we do not raise questions about ethical limits in their use, but rather we delve deeper into the nature of human relationships with other humans and with what we consider the techno-environment.

Ethical concerns come after metaphysical questions (ontological, epistemological and, above all, anthropological) and are better equipped to help us understand what a human being is and what the good means for them.[21] Given that technologies themselves – including anthropotechnics – also incorporate more or less implicit metaphysical assumptions, we could not tackle the moral challenges they raise without previously criticizing the metaphysical assumptions to which they respond.

Sometimes it is thought that an ethics without metaphysics would favour social consensus. But the truth is that such an ethics has never been found and, in any case, the loss of the

[20] M. Bertolaso and M. Rocchi, '*Specifically Human: Human Work and Care in the Age of Machines*' (2020) *Business Ethics: A European Review* (11 April 2020).

[21] A. Marcos, 'New Praxis. Releasement in a Technological World' in: L. Valera and J. C. Castilla (eds.) *Global Changes. Ethics, Politics and Environment in the Contemporary Technological World* (Dordrecht: Springer, 2020), pp. 17–27.

metaphysical horizon has only increased moral dissent and a lack of understanding among our fellow citizens. As we have said, we are dealing with two different anthropotechnical projects, with divergent objectives and disparate philosophical bases. The first one has defined as its goal to overcome the human. The second one proposes instead to apply technologies to improve or enhance human life, to make it more properly human. Here 'properly' refers to the very notion of 'human action' and the possibilities we always have to care for others and for common or relational goods. This is why, in our opinion, many discussions presented in this volume merit attention in relation to our ~~discussion~~ and further development to allow for a deeper understanding of basic human rights. For example, Douglas highlights how major difficulties in judging new technologies result from the difficulty in determining whether such technologies actually pose new questions or not with respect to preceding and more traditional procedures that did not necessarily make use of technology. It is not a matter of technology but more fundamentally of intentionality and respect for human beings in practice regardless of the level of technological mediation involved.

Only in this way will it be possible to develop and articulate further 'rights' and 'responsibilities', to justify an evolution of the debate about something as the 'human rights' that constituted for decades a pillar of our most developed societies and cultures. The challenge is to strengthen the roots and not to stress the multiplication of 'rights'. Through future technological developments, we may very well come to appreciate certain aspects of our human nature that were previously disregarded as superfluous, reductive, philosophical assumptions. Think of our mutual dependence and of the importance of solidarity or, as Botrugno discusses, 'face-to-face interaction must remain the gold standard of healthcare' and like parental care for children will always go beyond mere gatekeeper roles, assuring caretaker contextual conditions which are so crucial for a normal development for humans, as various recent studies also show.

Finally, as the European Group on Ethics in Science and New Technologies stated during the COVID-19 pandemic that has been affecting people's wellbeing and mental health, the pandemic has been a challenge to individualistic notions of safety, health and wellbeing, as well as for transparency, respect for democratic principles and accountability, asking for a strengthening of our values through a wider and more comprehensive understanding and application of solidarity principles.

The emerging notion of *technological humanism* amounts to all this. We should be able to make use of technology to guarantee those circumstances that really improve human life both individually and collectively and that permanently protect rights and liberties. Such a notion of technological humanism is clearly ~~mediated by~~ a specific understanding of human relationships that can obviously be mediated by technology but ontologically not substituted by it.