Technological Humanism & Sustainable A.I.

Alfredo Marcos and Marta Bertolaso

Abstract In this chapter we first disentangle the philosophical categories that are at stake when talking about 'digital environments' having in mind the question about sustainable AI. As ethics follows ontology, in the second part, we continue addressing some ontological questions on artificial intelligence (AI). The most elementary is whether the AI exists. We will argue that what actually exists is not AI as such but *AI systems* that emerge from the combination and dynamic interactions of human beings and machines. In this sense also environments that are digitally mediated deserve attention, philosophical and scientific analysis. In other words, we must go from considering AI systems as technical systems with social consequences to considering them as technically implemented social systems. People are part of the AI systems, as designers and users. As a consequence, we can say that in AI systems the intelligent part is not artificial, and the artificial part is not intelligent. The question about sustainable AI is thus not a technical problem, but an ontological and anthropological one, a problem of human ecology. When we accept that intelligence can be found in a simple device, we succumb to what might be called the *Toy Story effect*. Toys do not play by themselves, as well as machines do not have intelligence. Both playing and understanding require the concurrence of a human being, and this fact has a number of ethical consequences. Given these reflections, we conclude that considering the nature and the virtues that human beings can develop in environments that are also digitally mediated, opens new scenarios for education, policies and innovation that has to be sustainable by design.

A. Marcos (✉)
University of Valladolid, Valladolid, Spain
e-mail: amarcos@uva.es

M. Bertolaso
Università Campus Bio-Medico di Roma, Selcetta, Italy
e-mail: m.bertolaso@unicampus.it

A. Fabris, S. Belardinelli (eds.), *Digital Environments and Human Relations*, Human Perspectives in Health Sciences and Technology 150,
https://doi.org/10.1007/978-3-031-76961-0_10

Keywords AI · Delegated control systems (DeCo systems) · Information · Intentionality · Simulation · Function · Human ecology · Sustainability · Environments · Virtues

1 Introduction

Today we live in multiple environments. Naively, this notion denotes multiple spaces and contexts, such as, for example, the space inhabited by a person or the space that, generically, surrounds a thing but also, more broadly, the physical-chemical and biological conditions in which all living being live. In a figurative sense, the term can also represent the social and cultural dynamics that determine personal growth; in this second meaning of the term, a conceptual reference to the human being and his agency dimension in the world is still evident.

The notion of environment, therefore, semantically refers to multiple nuances and uses. Moreover, if we consider the technological development, we have to acknowledge that we live in environments that are more and more digitalized. Such transitions are bringing with them a new language such as the use of the term 'digital environments'. How should we understand such emerging concept? Should 'digital environments' be considered an additional 'environment' at the same level and with the same nature of the natural ones or should we deepen what is really at stake in the process of adopting such language to better understand the emergent risks and possibilities of the emergent technologies for an equilibrated human flourishing in the so called 'digital environments'? What epistemological and ontological issues are at stake? How can we reflect on them for an adequate topology of the environment notions and of the personal and social responsibilities we have to encourage and foster?

In order to answer these questions, in this chapter we first disentangle some philosophical categories that are at stake when talking about 'digital environments'. Secondly, as ethics follows ontology, we continue addressing some ontological questions on artificial intelligence (AI). We will argue that AI does not exist as such; what does exist, instead, are the *AI systems* that emerge from the combination and dynamic interactions of human beings and machines. In other words, we must go from considering AI systems as technical systems with social consequences to considering them as technically implemented social systems. People are part of the AI systems, as designers or users.

In the third part, therefore, the question about sustainable AI poses not a technical problem, but an ontological and epistemological one. Part of the argument follows also the consideration that when we accept that intelligence can be found in a simple device, we succumb to what might be called the *Toy Story effect*. Toys do not play by themselves, as well as machines do not have intelligence. Both playing and understanding require the concurrence of a human being, and this fact has a number of ethical consequences. Something similar happens with the notion of metaverse. Therefore, we conclude that human beings' flourishing and virtues have to be

developed in also digitally mediated environments, and that this requires new awareness and it opens new scenarios for education, policies and innovation that have to be sustainable by design.

2 Environments and Human Living Beings: An Integral Ecological Approach

Philosophically, we can start from the incontrovertible fact that the notion of environment is a concept full of meaning that refers to the age-old question of the relationship between context and living being and the mutual influence and relationship between the two. Different approaches can be adopted when trying to scientifically address this issue. Following Jakob von Uexküll and the notion of environment (*Umwelt*) he developed in his biological studies and in his book entitled "Animal Environments and Human Environments" (Uexkull, 2010; firstly published in German in 1933), it is to some extent commonly now acknowledged that the living beings might be considered as a zero point of a world or environment which is not unique and universal but constituted by and for the subject itself. The environment, therefore, is given to the subject in a changing and reciprocal exchange. The use of the plural—environments—is consistent with this idea. What also follows is that the environment cannot be enquired in isolation and that a physiological investigation that considers any living being as an object of research in itself, located in a unique and univocal world and that can be investigated in a universal way through a mechanistic approach, does not hold. At least, such reductionist approach is not useful when the question is about the sustainable development of a given living (eco)system.

In this sense, the theoretical revolution exemplified by Jakob von Uexküll mainly concerns the abandonment of the mechanistic perspective and the consequent reductionism represented by it in the ecological sciences. In doing so, his approach has contributed to dismantling the idea of the clock, made of mechanisms and gears, as the paradigmatic symbol par excellence of the constitution of the world and of the human being himself. Many works followed which recognized the notion of system and environment as constitutively relevant, undermining the modern reductionist views.

When coming to current digital transitions, the panorama seems to be more complex. If it is true that we live in physically characterized environments, we should ask the question whether the pervasiveness of the new digital tools is actually bringing us into a new kind of environments—"digital environment"—or what is actually at stake in these transitions in philosophical terms. We surf online from morning to night, we are always hyper-connected through social media, emails and instant messaging apps which have revolutionized communication. Chat-GPT is not an exception. We have the possibility now to develop economic transaction and market activities in contexts that are almost completely digitally mediated. Compared to the environments we were used to live, the digital imposes itself with a disruptive force,

in the creation of a non-physical but nevertheless immersive environment that can modify our experiences. Artificial Intelligence (from now AI) is the technology behind all these scenarios, it amplifies our possibilities of choice, of knowledge organization, accelerates our decision-making processes while orienting them in many cases. In order to analyse this new way of living, we will delve into three ideas: the ecological approach to the environment, the notion of environment-world and the revolution imposed by the digital transitions.

2.1 *"In-Being" as a Descriptive Character of the Notion of Environment*

From the description of the various environments made above, it emerges that the notion of environment thematically refers to the notion of living which, in turn, necessarily opens up to the methodological question of how we live in the environment. Parallel to the conceptual development of the scientific notion of environment, which we historically date back to the last century, the problem of finding a description capable of emphasizing the relationship between a human living being and the surrounding environment emerged.

On the basis of the made considerations, the relationship between subject and environment can be usefully described according to the Heideggerian philosophical category of "in-being" (Heidegger, 2008). This expression, taken from the existential analytics of the German phenomenologist, seems to capture the vital and relational nuance that emerges from the common conception of the environment as a place of experience that the subject inhabits. In fact, the notion of "in-being" (*Da-sein*) does not have the objective of describing the physically connoted character of an object of being inside something, as in the case of the water that is found in the bottle or the wardrobe that is in a room. On the contrary, such category underlines the existential relationship that is established with the world, that is, inhabiting the surrounding environment according to the character of familiarity. In the text *Fundamental Concepts of Metaphysics* (Heidegger, 2001), Heidegger deepens and further refines the distinction between world and environment by describing a human world-environment in which the human being, by nature, structures the surrounding environment as a world, making it not only the place of action and experience but forging and transforming it as a historical-cultural space. Understood in this way, the environment of the human being is an open world-environment, flexible and indeterminate in principle but determinable through the action of the human being himself. The transformative action of the human being, therefore, produces an always mediated world. Technology is a natural and constitutive part of such mediation. As we have many times argued, human beings are in fact naturally technical beings, i.e. they dwell in their environments and change or adapt it to their needs and expectations through the *techné*, a notion embedded in the modern technologies too. In this sense, we consider the environment as 'natural' (whether natural, i.e.

physical, or artificial). Artifacts are natural constituents of the human environment. But it seems clear that the digital mediation of some aspects of human environments poses new questions, as they change the perception and the possibility of action that human beings have in a given context. Considering the reciprocal interactions between the subject and the environments, this has to have consequences for the perception of the *Self* too.

2.2 Technological Mediality

Keeping in mind our initial question about sustainable AI, let's disentangle this point by considering that *mediation* is the fundamental fact of the technique and, therefore, of the technology with which the human being inhabits the world-environment and at the same time the characteristic that brings the concept of environment back to the centre of the contemporary debate. The advent of digital technologies has actually increased such level of mediality in human life and intersect previously isolated media fields, challenging the human "in-being" in a given environment.

What does this mean? As we discussed in another publication (Capone et al., 2023) when considering the impact of AI in the current digital transitions we have to move beyond a mere instrumental conception of the media, that is, thinking that "the purpose of digital mediation is merely the solution of a communication problem: representing any kind of content, conveying messages at high speed, at a minimal expenditure of information and at low cost (Shannon & Weaver, 1963)" (quoted in *ibidem*). The digital cannot be considered as a mere encoding as it does not merely deal with *the way* people solve a mathematical problem, listen to music or go shopping. "Media structure a model of relationship with things and these relationships are bearers of peculiar kinds of agency" (*ibidem*). Therefore, the human world-environment progressively takes shape as a mixture of different relationships, variously connected, each subjected to different mediations and liable to come into contact with certain media and not with others. Moreover, the media (be they platforms, particular technologies, information exchange protocols, methods of accessing data, etc.) direct human practices in certain directions and shape relationships within these practices.

2.3 Human Ecology and Technological Sustainability

According to Lady Bird Johnson's definition "*The environment is where we all meet, where we all have a mutual interest; it is the one thing all of us share* " (Johnson, 1967). This concept is the basis of ecological philosophy, which examines humanity through an eco-ontological lens, highlighting relationships and sharing (Marchesini, 2002). This perspective recognizes that everything in the world is

interconnected, living beings and the environment are intrinsically linked, interdependent. Humanity is an integral part of the ecosystem, it is not separated from it, and every human action, every economic or social process has an impact on the environment and on human beings themselves. The environment is one of the fundamental dimensions of this interconnection, influencing and being influenced by every aspect of human life and the natural world. The above-mentioned environments—also technologically mediated—do not exist in the vacuum but have consequences for all the other living beings' environments and vice versa. It therefore makes sense to think of these issues in terms of *Human Ecology* to remind us that reflection on human behavior cannot be separated from the awareness of the interconnections that exist between us and the world around us. Somehow the relationship we have with ourselves, the care we are able to take of ourselves and others necessarily reflects our relationship with the world and the world we want to build and inhabit.

Philosophers such as Aldo Leopold and Arne Naess have developed innovative ethical paradigms that redefine human-environment interaction on the basis of a notion of environment that is fundamentally linked to that of resource: that is, the resources that the environment offers and all living beings they have an intrinsic value, and this must push humanity to reconsider its role in the Earth's ecosystem and to develop a deep sense of responsibility towards the environment in which it lives and operates (Leopold, 2020; Naess, 2016). The environment, therefore, as a constitutive dimension of the dynamics of a system, requires people to have an ethically responsible approach that is aware of the ecosystem value of the environment. From this come the objectives of preserving natural resources, which would continue to support life on Earth, and that of promoting long-term global well-being without compromising the ability of future generations to meet their own needs—sustainable development [see Brundtland Report (UN, 1987)]. A healthy environment is essential to ensure that this equation is met, and is also capable of adapting to change, which is crucial for the resilience of today's rapidly evolving world. The global challenges posed by the goals for the sustainable development (SDGs) follow this logic.

Technology contributes to these processes. In particular, digital technologies can be integrated into physical environments to create advanced sensors that allow interaction between humans and the surrounding environment, providing real-time data (monitoring), which can be made accessible through applications or digital interfaces (implementation). In this sense, the digital technologies can also make environmental 'resources' more transparent and allow us to influence them like never before. Network logics and the dematerialisation of digitally mediated relationships allow us to manipulate and orient environmental dynamics by adapting them to individual (personalization) and collective (interactivity) needs, allowing users to examine specific aspects, such as (i) identify complex relationships between the individual and various other factors, (ii) formulate predictions using systemic logics and (iii) make more informed decisions about the environment and people's health too. In both cases, nevertheless, digital technologies create interfaces that are nevertheless *representation of the environment-world*. Although such representations

sometimes also end up being part of our environment, their representational nature what actually affects are the human relationships and their perception.

As in the case of technologies applied to the human body, also in this case we need to question ourselves about the constraints that hold a flourishing use of such technologies in mediating a perceived environment. As Hans Jonas said, losing the human for enhancement technologies would cause us to lose the axiological reference that allows us to evaluate its goodness or illicitness (also discussed in (Bertolaso & Marcos, 2023)). It thus makes sense asking the question of which axiological reference is necessary consider so as not to disorientate ourselves in incorporating technology into the environment around us. In this case we will talk about eco-technological sustainability, which deals with the still only human capability of making decisions. Beyond our possibility of action human intelligence is up to this: questions about meaning, and actions that follow a judgment. Values and virtues mediate this dual processes. A relational philosophical understanding of the human beings holds the possible integration of all these aspects.

3 What Is A.I.

After having contextualized the environment issues and highlights reasons for an integration of the philosophical questions at stake across the notions of environment, digital technologies and sustainability, the need for an ontological reflection emerges. As ethics follows ontology, the question about A.I. (what is AI) still needs to be addressed in the light of a philosophical account of the human beings as relational and technological beings, beings that are able to change the environment for their own sake and objectives.

Heidegger "in-being" notion requires, therefore, not only a deeper reflection about the human beings, but also about digital mediated environments and AI more in general.

Although we all have been reading and writing a lot about AI, we suggest now to consider AI ethical and political problems from a different perspective. Our ethical tribulations, in fact, almost always refer to deeper ontological problems that sometimes we have undermined. And so is the case regarding AI. Without addressing the ontological questions, the ethical debate becomes imprecise and superficial. Disentangling the ontological questions that are behind the ethics of the so-called AI, we find that the most elementary question we can ask in ontological terms is whether or not AI exists. Luc Julia (Julia, 2019), who led the SIRI development team at Apple, provocatively titles his book: *L'intelligence artificielle n'existe pas*. For his part, Erik J. Larson (Larson, 2021), in a similar vein, has published the book entitled *The Myth of Artificial Intelligence. Why Computers Can't Think the Way We Do*. In the advance of this book we can read: "We aren't really on the path to developing intelligent machines. In fact, we don't even know where that path might be […] AI will continue to improve at narrow tasks, but if we want to make real

progress, we will need to start by more fully appreciating the only true intelligence we know—our own".

3.1 An Inappropriate Name

What do we mean when we say that there is no AI? Maybe it's (again) a naming problem. Perhaps there is something which we call AI, but for which we are using an inappropriate name, since it does not contain true intelligence or the intelligence it contains does not reside in the artificial part of this entity. It would, therefore, be necessary to propose a better denomination, which would not lead to confusion. "The term artificial intelligence – recalls Katharina Zweig, from the Algorithmic Accountability Lab at the University of Kaiserslautern – arose in the 1950s, when scientists wanted to raise money for their research. They thought it sounded like something the State would be happy to encourage. And now we hang on this name. Most computer scientists find it inappropriate" (Von Hopffgarten, 2021).

Machine learning or *deep learning* are equally confusing names, especially if we have to understand that it is the machine itself that learns. All these denominations—lures, it should be said—have a commercial, advertising, even propaganda function, but they do not respond to the truth of the thing. They immediately resonate with science fiction and the media headlines. Therefore, dreams and future scope terrors begin to thrive. However, no machine understands, or knows, or learns, or is capable of counting to two. People do it, with the help, sometimes, of machines. For this reason, other terms have been proposed, such as *assisted intelligence*, *expanded intelligence*, *human-centered artificial intelligence*, *decision support tools*… These names are more appropriate, since they indicate that the intelligent subject is a person, while the machine can assist or expand the intelligence of said subject. We could also talk, and I think it would be the most appropriate option, of ***Delegated Control systems*** (which we can abbreviate as *DeCo*).

The change in perspective that this name introduce could be summed up in a few words: we must move from considering AI systems as "technical systems with social [external] consequences" to considering them as "technically implemented social systems" (Hirschheim et al., 1995 [1]). In other words, people are also part of the AI systems, as designers, owners, maintainers, users, supervisors, lawmakers… It is in these people, and not in the artificial part, where the intelligence of these systems resides. In AI systems, the intelligent part is not artificial, and the artificial part is not intelligent. Machines cannot be intelligent by themselves, outside a human environment. This limitation does not respond to a technical problem that can be technically corrected, but to an ontological difference.

3.2 So, What Is Intelligence?

Now let we take a step back and reflect for a moment on the very concept of intelligence. This way, we will better estimate if certain entities deserve or not to be called intelligent. Or, more precisely, where intelligence resides in *a technically implemented social system*. Dictionary definitions of the word "intelligence" often refer to the ability to understand. It is also common for them to allude to the ability to solve problems. For instance, The Oxford Learner's Dictionary defines "intelligence" as "the ability to learn, understand and think in a logical way about things". For the *Diccionario de la lengua Española (Spanish Dictionary*), intelligence is the "ability to understand or comprehend… [and, in a second meaning] The ability to solve problems". And something similar appears in the dictionary of the Italian language *Treccani*.

We know that the artificial part of an AI system is incapable of understanding by itself, without humans. You cannot even properly say that a machine counts or computes. Counting implies joining two (or more) moments (or things) and keeping them together in a single and identical conscious representation, understanding at the same time the similarity—not identity—and the difference between them. All this far exceeds the capabilities of a machine.

It is true, on the other hand, that AI can help us solve multiple problems (computing, writing texts, drawing, geolocation, logistics, telephone assistance, medical diagnosis assistance, advertising and a long etc.). But these problems are not for the artificial part of the system, but for the human designer or user of the system. For a facial recognition machine, recognizing or not recognizing a criminal is not a problem. It is a problem for people's safety, and the system can help us deal with it. Of course, the same system can be used to control the population of a country and to facilitate political repression there. Again, this is not a problem for the cameras or for the software involved. It is, undoubtedly, a problem for the human subjects of the country in question. Only a living being can suffer and die; only a person can wonder about the meaning of her life. Those are problems. And both a hammer or an abacus, and a computer network too, each in its own way, can help us deal with these problems (or make them worse). But this does not make these tools intelligent. We understand. We have problems. Not the machines.

Let's take a look from another angle. Sometimes the so-called AI is characterized by its simulation capacity. It simulates functions of human intelligence, it is said. However, simulating intelligence is not the same as being intelligent. Furthermore, the simulation only appears as such for the human being who observes it, not for the machine. The machine does not know that it is simulating intelligent behaviour. On the other hand, the very notion of function inexorably refers to a being for which a given effect is functional. Outside the human framework, the lights that come on and off on a screen, or the movements of a robot are mere effects. They do not fulfil functions. It is the human point of view that changes its ontology transforming mere effects into functions.

3.3 Back to the Ontology

So, the question is basically—let us insist—of ontological nature. Artefacts, in the Aristotelian tradition, in the Aristotelian ontology, are substances only in an accidental sense, i.e. by analogy. Living beings, and especially human beings, are substances in their proper and paradigmatic sense. As it is an ontological difference, the hope (or threat) of annulling it through technological sophistication is illusory, a mere category mistake.

What has been said so far affects any technological system (washing, transport, energy production, communication…). All of them, if they are placed outside of the human environment, lose their functionality, they become plain systems of physical effects. Since their ontology is given by their function, they also lose their ontological rank, they stop being what they were. A washing machine placed on Neptune is no longer a washing machine. Nevertheless, the ontology of the so-called AI systems depends even more intensely on the human gaze, since they are located in the realm of the intentional, that is, of the semiotic. In this area, the entities are sustained on three supports. If we remove one of them, they will collapse, as happens to the stools. Charles S. Peirce makes it clear: "All dynamical action, or action of brute force […] takes place between two subjects […] But by semiosis I mean, on the contrary, 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" (Peirce, 1935 [484]).

What do we mean when we say that a machine stores or processes my financial or medical data? We say that a certain electromagnetic state of the machine (*sign*) is related to my payroll or my blood pressure (*object*). Obviously, there is no physical relationship between them, but rather a semiotic relationship that is established through a person (*interpretant*) capable of understanding or interpret—with the help of certain interfaces—electromagnetic states as income or blood pressure. Similarly, the machine only plays chess or *go* if a person can relate the physical states of the machine to these games traditionally played by humans. The case of chess is very illustrative: when a machine finally meets certain expectations, it is at the same time deprived of the mythical aura that surrounded it when it was just a project; it is reduced to the level of the prosaic, devoid of ghost and glamour. Look at the poor *Deep Blue*, who knew glory days, raising now museum dust.

It is exactly the same in the case of the now famous ChatGPT. Its performances are really impressive. It is very helpful in multiple tasks. But, in spite of its astonishing attainments, no one thinks that this software understand anything. Rather, it has become clear to us that great linguistic achievements can be attained without understanding anything at all (as John Searle advanced years ago).

So, we've all learned, at the end of the day, that a chess-playing robot, or a chatbot, is about as interesting as a vacuum-cleaning robot. Without an *interpretant*, the machine just changes from one physical state to another. It is no longer part of an intelligent system. It is just a piece of matter, like a washing machine on Neptune. Why is it so hard for us to accept it?

4 The *Toy Story* Effect

We tend to imagine that in our absence the artefacts continue to have the same entity as in our presence. Thus, we imagine that a machine that is part of an AI system, together with certain people, continues to be intelligent even if it does without the gaze of those people. But this mirage is not due to an excess, but to a lack of imagination. It is not easy to imagine what the world looks like when the world is not seen by anyone. The look of the human being sustains the being of the artificial. Without the gaze of a person, the artificial is flattened; it becomes in pure physical reality. Hence the difficulty of imagining. It's easier to dream that everything stays the same when I stop looking. It is what we could call *the Toy Story effect*. The child's hand and eye turn a piece of green plastic into a shy dinosaur. The child imagines that when he leaves the room, the dinosaur is still there. He can't think of it like the inert piece of green plastic it is when the child himself leaves the playroom or falls asleep. ("When he woke up, the dinosaur was still there". This is a famous flash fiction by Augusto Monterroso (1959)).

What do we imagine would happen if the human beings left the room (Chinese?), if they were left out of the AI systems? For some, this will occur from the point they call *singularity*. From there the machines would generate other smarter machines, a post-human world controlled by robots. But perhaps we could imagine, on the contrary, that machines left to themselves would soon fail by virtue of the general tendency towards entropy, design and construction defects, as well as the difficulty in obtaining stable energy sources; they would decay and be re-incorporated into the natural world, into the physical processes (such as erosion) and chemical processes (such as oxidation). Even into the biological processes: the most probable post-human landscape is not that of the Earth governed by intelligent robots, but that of a leafy jungle that hides in its bowels, along with the stones of ancient temples, authentic filth of silicon, plastic and metal. In fact, every machine has to be maintained, that is, led by the hand by people ("hand" in Latin "manus", and from there, in English "to maintain"). Every AI system requires maintenance. And the more sophisticated ones require more maintenance, not less.

In short, data are data about something, intelligence is intelligence about something, and also the information. They are triadic, semiotic, intentional entities. The electromagnetic (or quantum) states of a computer are not data by themselves, unless an *interpretant* manages to connect it with an *object*. (And this dependence of the data on a consciousness is extreme in the case of the so-called *synthetic data*, whose relationship with the truth is so difficult to establish). Without people, an AI system immediately ceases to be intelligent. There are no more data. It no longer understands anything. It no longer simulates anything. It no longer serves any function. Its troubles are finally over. What we called information is diluted. All decision ceases to be such.

In order to avoid misunderstandings, we should change "AI" to a less tortuous name, since the disturbance generated by a bad name ends up being projected onto anthropology itself. Thus, the image of human intelligence is degraded or reduced

to an algorithmic game. Such a game could be played on this or that material support (carbon, silicon… it doesn't matter). So, the human being himself comes to be understood in dualistic terms, as a kind of fortuitous meeting place between a bodily hardware and a mental software susceptible of migration to "the cloud".

4.1 Metaverse: What Are We Talking About?

Having established the renewed status of the environment, we can say something more about the most paradigmatic—at the moment—example of the so-called 'metaverse'. As discussed, we often inhabit non-extensive places and we depend on them in many dimensions of our existence; from work to social life, from family to public relations, everything passes through the filter, for example, of the internet or digital platforms. The smartphone has become, for all of us, an extension of the arm and often also of our mind. Even more extreme situations occur with augmented reality and virtual reality, which insert a constant and continuous interaction between the dimension of reality and virtual parallel worlds.

In this sense, the metaverse seems to represent a *liberation from space* which has passed through the de-structuring of activities previously located within rigid perimeters based on the Aristotelian units of space-time-action. All this leads us to ask again with (Heidegger, 2001): what does space become today and what are the capabilities necessary to inhabit it and, therefore, to build it? In redefining the idea of space, technology and the metaverse play a crucial role as reality is aimed elsewhere, no longer located. But in any form, it substitutes what we know as 'environment'. The metaverse, in fact, while apparently healing the classic Cartesian body-mind/hardware-software dualism, as a simulation of some environment's features, cannot be considered an alternative to it. Metaverse's value relies upon the kind of relationships it is able to represent and leaves many open questions regarding its ability to make us experience the environment and human relationships. As it stands, it seems to be more extractive that enabling of human relationships, with other or with the real environment. The difference is not, however, in the available technology but in the virtues, habits, social practices and human expectations that push people to use the metaverse technology.

Consistently with what we have said above, without people an AI systems and virtual contexts immediately stop working, AI no longer 'understands' or even 'simulate' anything. No virtual reality, no *metaverse*—and therefore no *digital environments*—actually exists by themselves either. Again, a digital representation of reality exists—of the only reality that exists—within which the virtual has a place as a representation, but decays when there is no one to represent something to. In this sense -we agree with Luc Julia-, AI does not exist. It is the mere effect of a poor imagination, that can unfortunately be placed at the service of economic or political ambition.

The barriers that the metaverse is therefore helping to overcome is not space and time as such, but the strictly link we previously experienced of a one-to-one

relationship between function and space. In doing so, the conventional closed and delimited space can evolve into an open, multifunctional space, intended for a fluid and plural use, which aims to define new places and new ways of inhabiting a post-topic world-environment while always maintaining the existential structure of human relationality. Functional goals a therefore still mediated by the AI technology, not question about meaning, opportunity, relevance, justice.

5 Sustainable AI

"Bad theories destroy good practices," (Ghoshal, 2005) said. The world of the digital innovation is no exception. We are seeing things that work (from digital solutions to the metaverse, from open spaces to phygital meetings), but the reason for their effectiveness and above all the conditions of sustainability of these innovation processes may not hold in virtue of such technological advancements. As said, supporting them without reflecting on the epistemological and axiological foundations means running into the danger denounced by Ghoshal and which is already a reality, for example, in the seven sins attributed to Greenwashing (CIT…): the sin of omission of information, the sin of no proof, the sin of vagueness, the sin of irrelevance, the sin of the lesser evil, the sin of lying, of adopting false labels. In this sense, a superficial technological innovation is possible, as a superficial ecologism is.

A sustainable AI technology development is possible on the basis of human choices and responsible technology's use. The pervasiveness of the digital tools asks for a deeper reflection about human relationships and desire. A different model of 'progress' is fuelled with narratives that do perpetuate wrong socio-imaginaries and toy stories as discusses. The idea that machines can at some point substitute human relationships. We can choose surrogate of them: this will be the problem, not that we can generate surrogates. We do so all the time, from food to clothes and even entertainment and art.

Narratives and toy stories affect human beings' expectations and prevent a serious process of personal and collective awareness and responsibility about what is at stake in the digital transitions. An innovation projected towards growth that reiterates productive and technological innovation like those that created the problems we suffer, systematically leads to avoiding reflection on limits and boundaries, which are instead generative of new practices and new economies. Even when we emphasize the importance of relationships and interdependencies, it seems easy to forget that taking them seriously means adapting, changing pace, dealing systematically with mutual dependencies where what actually make the difference are the human values, real relationships, human decisions and judgments.

In this sense, innovation is sustainable as far as it aims at the human flourishing, or it is not innovation at all. Its sustainability grounds in the ontology of the human beings, of the real world as our given environment and of AI as a tool. Such sustainability can be pursued as far as we take care of the motivations that generates, enhance AI technologies and encourage its use, not through a mere calculation of

risks and of possible impacts. It is a social and educational problem, not a technical one. Again, it is the result of choices not of necessity.

Going back to the notions of environment and human ecology, the anthropological reason of what we are saying here is that any true innovation has its objective in the caring of other and of their place (environments). Places that should be understood as spatio-temporal experiences of co-existence with us ourselves, with others and with the world we inhabit. We have an obligation to care, to take care of ourselves, of others and of the ecological issues. It is a *human ecology* challenge.

For true innovation it is necessary to develop a new reasoning on human duties (even first than on human rights). Such duties exist in everyone's conscience even if no one were to recognize them, Simon Weil (Weil, 2017) would say. For sustainable innovation, we do not need regulations in the first place, but a new personal and collective moral and ethical conscience.

5.1 Prudential Intelligence

Moreover, when dealing with *an uncertain world*, we have to use not only the brute force of algorithms, but also the imagination, creativity, intuition and prudence human beings are capable of. In fact, the most comprehensive "method", the one that regulates the application of all the others, including the automatable ones, is human prudence.

As far as we know, the universe is not some kind of eternal clock, but a unique, historical, and contingent event. It is endowed with a certain network of regularities, sufficient to make life and intellection possible, but compensated with unpredictable novelties. This very peculiar distribution of constancy and rupture affects both the orbit of the planets, as well as our daily life, made of imperfect cycles, circadian rhythms, habit and shock. Only a living, sentient, located and interactive intelligence, a prudent intelligence like the human, can understand this disconcerting texture of the universe. We learn from experience, of course, but we know at the same time that there is no guarantee that things will continue as they were. Hence the convenience of intellectual humility, which has been dressed over time as a Socratic attitude, Aristotelian prudence, learned ignorance, fallibilism…

An AI system generates expectations. (It places a point in an n-dimensional space constructed from a history of data, and, based on it, it tells us what can be expected regarding the object represented by that point.) But the system can crash when it registers the occurrence of something whose possibility was not even considered in advance. When this happens, the system itself is left without the ability to adapt, it cannot learn from this experience. When this happens, it is not the algorithms that must react, but the people responsible for them. And they will react, first of all, by drastically changing expectations. They can do it since they are not artefacts, but conscious people who can come to understand the phenomenon which is not expected from the machine. Humans can activate their creativity to generate better expectations from now on with or without mechanical help. A person can

conjecture causal relationships, beyond the mere correlations that a machine detects. And this step does not have to be purely arbitrary, random or irrational, but rather, in some sense, it is guided by a practical and social knowledge that Aristotle called *phrónesis*, prudence. Said knowledge facilitates (*i*) the integrative constitution of the experience, (*ii*) the management of emotions linked to the frustration of expectations, (*iii*) the propaedeutics of the creative moment and (*iv*) the critical filtering of the new emerging expectations.

In the words of Erik Larson: "AI works on inductive reasoning, crunching data sets to predict outcomes. But humans don't correlate data sets: we make conjectures informed by context and experience. Human intelligence is a web of best guesses, given what we know about the world. We haven't a clue how to program this kind of intuitive reasoning, known as abduction [(Aliseda, 2010)]. Yet it is the heart of common sense [prudence]. That's why Alexa [or SIRI, or ChatGPT] can't understand what you are asking".

Consequently, an AI system that aims to replace human prudence would simply be out of place, out of the universe that houses us (out of our uncertain world). On the contrary, a DeCo system inscribed within the framework of the prudential human intelligence will be in its rightful place and will be able to fulfil functions of a great value for human life.

6 Final Practical Orientations

The practical problems have, thus, nothing to do with a supposed post-human future of intelligent machines. "What should terrify us -says Ramón López de Mántaras, founder of the CSIC's AI Research Institute – is not a future dominated by a hypothetical superior AI […] What should really worry us is the present situation, in which we are delegating more and more tasks in an AI as limited as the current one" (López de Mántaras, 2020 [59]). And "delegating" is the key word here. What is crucial has to do with the present, with the way in which DeCo systems are already being used, with the responsibility that certain people, companies and governments have for it, as well as with the impact that this use already has on our life.

The new name (DeCo systems (Bertolaso & Marcos, 2023)) does justice to the true ontology of these systems and, above all, leads us more directly to important practical issues. We begin to see what the relative position of humans and machines should be. It is not a matter of deforming the former, the humans, to fit into a world presumably dominated by mechanical intelligences, but of placing the latter, the machines, within the framework of human life. Outside this framework, as we said, they cease to function, even cease to be what they were. What algorithms can contribute to human life? They allow us to delegate the control of certain processes.

Only at this point we can more appropriately raise the pertinent practical issues, ethical, political, educational, legal… Whose is the hand that rocks the algorithms? Who delegates? Are they entitled to do so? In which DeCo systems do they delegate? Are they the most appropriate systems? What kind of actions/processes are

delegated? Are they really delegable?, technically?, ethically? For how long is control delegated? Is it sensible? Is the delegation reversible? What monitoring or evaluation procedures exist? Are they sufficient? What risks are assumed in case of failure? Is it prudent to assume them? What advantages are obtained for human life with the delegation of control? What do we lose in return? …

At the educational level, it is, of course, essential to teach young people (and not so young) certain values and train them in virtue, rather than insist on instructing machines in the ideology of political correctness. Unfair biases will not be redirected just by redesigning algorithms, but by educating virtuous people. However, virtues and values have to be educated to be at the level of the current technological context, as a part of a new *technological humanism*. For example, through intermittent technological silence practices (Marcos, 2020).

At the political level, the DeCo approach allows us to immediately identify the legitimacy deficit and the risk to people's freedom. Thus, most of the DeCo technologies are ultimately in the hands of the Chinese Communist Party and a few large US corporations. The greats of the so-called AI are, on the Western side, the GAFAM (Google, Amazon, Facebook, Apple and Microsoft) and, on the Chinese side, the BATX (Baidu, Alibaba, Tencent, Xiaomi, and you can add now Tik Tok or Huawei). The first impression we get is that there is too much power in too few hands.

We have to ask ourselves if the power they hold is legitimate. If it is not, neither its delegation to algorithms will be. It is imperative, for the health of democracies and for the freedom of the people, that the excessive power of these corporations be dissolved. Responsible digital consumption, lucid use of social networks and prudent management of each one's own data would already help (Véliz, 2020). But in addition, political pressure and public opinion in favour of the dissolution of nuclei of abusive digital power would be pertinent.

Just one example of something that is happening right now, in the last few days. Google has recently developed a tool, *Genesis*, designed specifically for the media, for the preparation of news for the media. They have already presented this application to the main North American media, *New York Times*, *Wall Street Journal*, and *Washington Post*. Through this software, Google could obtain a determining influence in the construction of world public opinion. We cannot leave the construction of the media vocabulary and the media agenda simply in the hands of Google. We don't want *Genesis* to become *Apocalypse*. At least, it is essential to demand a public clear labelling of the media pieces built by an AI system.

The new European AI law already takes steps in the right direction. The USA, UK and the G-7 have also taken steps, in recent days, towards the legal control of AI. Even the UN has recently issued a document with recommendations on the governance of AI.[1] However, as AI is now in the hands of half a dozen US corporations and the Chinese Communist Party, a determined activism in favour of the fragmentation of digital power is required. It must be a simultaneous task to the

[1] See https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf (accessed May 12, 2024).

proclamation of laws. Without addressing it, any restrictive European legislation will only serve to make Europe even more irrelevant in the AI concert. Perhaps also to relocate the most disturbing or dangerous applications of the AI to the most vulnerable countries.

Let's finally go back to the so-called digital rights. It should be clear by now that robots have no rights, nor are they responsible for anything, nor do they have to pay taxes, that it is absurd to talk about software rights. We are before a category mistake. But neither does it seem very perceptive or very useful to ask for a new generation of human rights, this time digital rights –let's say. Inventing new human rights weakens the very idea of "human rights". Human rights are based on the dignity of the person and their belonging to the human family. It is always tempting to use this formula –"human rights"—to protect any asset considered valuable. But, as the field to which this formula is applied expands, its protective force inexorably diminishes.

If we understand by human rights those of the first generation, fundamentally the right to life and liberty, then the accusation of having violated human rights is an extremely serious accusation. But if we include not only second and third generation rights, but also new digital rights and neurorights, plus the "human" rights of robots or animals, then the accusation of violation of a right becomes slightly disturbing. Much more insightful and useful would be to connect the digital to the basic human rights, to show how it affects them, without inventing new lists of supposed human rights. What is serious about some DeCo systems is not that they violate our alleged digital rights, it is that they can sometimes threaten our lives or compromise our freedom.

Acknowledgments We thank Laura Corti and Maria Laura Ilardo who are contributing to organizing some reflections of the first paragraphs of the current paper.

References

Aliseda, A. (2010). *Abductive reasoning*. Springer.
Bertolaso, M., & Marcos, A. (2023). *Umanesimo tecnologico*. Carocci.
Capone, L., Rocchi, M., & Bertolaso, M. (2023). *Rethinking "digital": A genealogical enquiry into the meaning of digital and its impact on individuals and society*. AI & Society.
Ghoshal, S. (2005). Bad management theories are destroying good management practices. *Academy of Management Learning & Education, 4*(1), 75–91.
Heidegger, M. (2001). *Fundamental concepts of metaphysics: World, finitude, solitude*. Indiana University Press. (new edition).
Heidegger, M. (2008). *Being and time*. Harper Collins. (reprint edition).
Hirschheim, R., Klein, H. K., & Lyytinen, K. (1995). *Information systems development and data modelling*. CUP.
Johnson, B. (1967, October 9). *Speech at Yale University, white house diary*. Retrieved September 8, 2024, from https://www.nps.gov/common/uploads/teachers/lessonplans/Lady%20Bird%20Redefines%20Beauty%20Curriculum1.pdf
Julia, L. (2019). *L'intelligence artificielle n'existe pas*. Éditions First.
Larson, E. J. (2021). *The myth of artificial intelligence. Why computers can't think the way we do*. Harvard University Press.

Leopold, A. (2020). *A Sand County almanac: And sketches here and there*. Oxford University Press. (illustrated edition).
López de Mántaras, R. (2020, July). El traje nuevo de la IA. *Investigación y Ciencia, 526*, 50–59.
Marchesini, R. (2002). *Post-human: verso nuovi modelli di esistenza*. Bollati Boringhieri.
Marcos, A. (2020). New praxis. Releasement in a technological world. In L. Valera & J. C. Castilla (Eds.), *Global changes. Ethics, politics and environment in the contemporary technological world* (pp. 17–27). Springer.
Naess, A. (2016). *Ecology of wisdom*. Penguin Classics.
Peirce, C. S. (1935). *Collected papers* (Vol. 5). Harvard University Press.
Shannon, C., & Weaver, W. (1963). *The mathematical theory of communication*. University of Illinois Press.
Uexkull, J. (2010). *A foray into the worlds of animals and humans: With a theory of meaning*. University of Minnesota Press. (illustrated edition).
United Nations (UN). (1987). *Report of the world commission on environment and development: Our common future. transmitted to the general assembly as an annex to document A/42/427–development and international co-operation: Environment*. Retrieved May 12, 2024, from https://www.unicas.it/media/2732719/Rapporto_Brundtland_1987.pdf
Véliz, C. (2020). *Privacy is power*. Transworld.
Von Hopffgarten, A. (2021). La inteligencia artificial carece de la flexibilidad de decisión humana. Interview with Katharina Zweig. *Mente y Cerebro*, enero/febrero.
Weil, S. (2017). *La prima radice* (F. Fortini, Trans.). Edizioni di Comunità.