

FOCUS 101 – Between digital revolution and anthropology: crisis of the Subject and hybrid societies

Pasquale Grieco

1. Vampires and Information Technologies

The Contemporaneity shaped by digital technologies is not just the product of an enormously increased computational power, which, as the Moore's Law¹ seems to suggest, grows exponentially. This age, our age, named by Luciano Floridi the age of the *Fourth Revolution*², is marked by a dual sense of change (and it cannot be otherwise, because it is its feature of *duality* that qualifies its revolutionary character): an extroverted change, directed towards the world³, and an introverted one, towards the individual.

¹ The Moore's Law was formulated by Gordon Moore, co-founder of Intel, and states that «The complexity of a microcircuit, measured for example by the number of transistors per chip, doubles every 18 months (thus quadrupling every 3 years)» (G. Moore, *Cramming more components onto integrated circuits*, «Electronics», 38, 8, April 19, 1965, pp. 114-117).

Far from being a “law of nature”, it represents a generalization indicating the exponential increase in computational power and the corresponding decrease in costs of the technology implementing it. Moore's Law has become the metaphysical basis for Ray Kurzweil's “faith”. Kurzweil is an entrepreneur, inventor, and Google's Chief Engineer, relentless preacher of the advent of Technological Singularity and the possibilities of Transhumanism. See R. Kurzweil, *The Singularity Is Near: When Humans Transcend Biology*, Viking, New York 2005.

² L. Floridi, *The Fourth Revolution. How the Infosphere is Reshaping Human Reality*, Oxford University Press, Oxford 2014.

³ Also noteworthy, even if I do not subscribing to them, are the physical, philosophical, and mathematical conceptions characterizing the entire universe as a gigantic computer (digital or quantum), whose structure and primal matter would be formed or derived from the Bit (*It from Bit*) (K. Zuse, K., *Calculating Space*, MIT Press, Cambridge Massachusetts 1970; J.A. Wheeler, *Information, Physics, Quantum: The Search For Links*, «Proceedings of the 3rd International Symposium on

These changes produced, and they are still producing, a re-ontologisation and a re-epistemologisation of our ideas about modernity (world, subject, thought, nature, culture, action, community, and so on) that are requiring effective critical tools to be adequately conceptualized.

The use of certain methodological results from cultural anthropology and historical studies can be fruitful in this regard. These outputs can be applied in topics deeply relevant to the historical-scientific, ethical, and epistemological transformations associated with the spreading of the *Information and Communication Technologies* (ICT) and its theoretical background.

Exempla like the phenomenon of shamanism, its strict connection with the unconscious structures of the mind, and the “automatic” or “mechanic” nature felt by everyone during simple tasks such as stroking keys on a QWERTY keyboard⁴, driving a car, or engaging in other

the Foundations of Quantum Mechanics», Tokyo 1989, 354-368; E. Fredkin, *An Introduction to Digital Philosophy*, in «International Journal of Theoretical Physics» volume 42, pp. 189-247). For a quantum interpretation of the universe as a computer see S. Lloyd *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos*, Knopf, New York 2006). In this framework, according to a certain “ontologizing” characterization of the Church-Turing Thesis (which states that everything computable is computable by a Turing machine) computers would have an effective capacity for creation/transformation, writing/rewriting of the Universe. These theses also underpin much of the Simulation theories, suggesting that we might be living in a gigantic virtual simulation, or at least lacking any conclusive proof to distinguish a “real” reality from a simulated one. See D. Chalmers, *The Matrix as Metaphysics*, *thematrix.com*; reprinted in Christopher Grau, ed., *Philosophers Explore the Matrix*, Oxford University Press, Oxford 2005, pp. 132-76, and D. Chalmers, *Reality+. Virtual Worlds And The Problems Of Philosophy*, W.W. Norton & Company, New York 2002; N. Bostrom, *Are we living in a computer simulation?*, in «Philosophical Quarterly», 53, 211, 2003, pp. 243-255. For a philosophical overview on the theme of the informational conception of physics see D. Chalmers, *Reality+*, cit., pp. 148-160). For a critique of the above-described digital ontology, in favor, instead, of a defence of an informational structural realism, see L. Floridi, *The Philosophy of Information*, Oxford University Press, Oxford 2011 pp. 316-328.

⁴ The QWERTY keyboard is a good example of automatic knowledge interiorized by the body. Touchtyping is an «incorporating practice», an action that is encoded into bodily memory by repeated performances until it becomes an habit. T. Bardini, *Bootstrapping: Douglas Engelbart, Coevolution, and the Origins of Personal Computing*, Stanford University Press, Stanford 2000, p. 67. Doug Engelbart, the

routine activities, are not connected merely by chance. These instances can be observed in the historical development of human-machine interactions.

It is not accidental either that there is a relationship between the typical shamanic alteration of states of consciousness, induced by drugs or other environmental causes, and the historical role played by LSD as a form of “liberation of the mind”—with the mind considered cybernetically as an information processing system—in the development of the contemporary shape of our personal devices and digital and Internet technologies⁵.

Among the traditional categories that the contemporaneity requires to put under critical lens, particularly urgent is analysis of the Subject. What we call Self, with its set of quintessential features (simplicity, self-transparency, rationality and so on), is maybe just a cultural construct of a particular period of the intellectual history: the Modernity. The digital revolution is a further step toward the crisis of this construct. Applying anthropological and historical analyses is helpful

man whose vision has shaped the contemporary personal human-computer interface attempted, instead, to develop tools capable of augmenting the automatic bodily (and mental) capabilities of the user. Cfr. Ivi, pp. 58-80.

⁵ In the development of augmenting interfaces, there is an instance red thread in the history of informational, cybernetic and computing thought that aims at the automatization of lower/automatic function of mind activity in order of freeing up higher-level ones. Cfr. Ivi, p. 19. Doug Engelbart, as many of the artificial intelligence and personal computing pioneers, experimented with LSD. S.A. Kallen, *The Information Revolution*, Lucent Books, Farmington Hills MI 2010, p. 32. On the role of the LSD and other psychedelic drugs, included peyote and hallucinogenic mushrooms used in shamanic rituals, in the biographical and intellectual events of computer revolution main protagonists, included Steve Jobs, see. J. Markoff, *What the Dormouse Said: How the Sixties Counterculture Shaped the Personal Computer Industry*, Penguin Books, New York 2005. The role of counterculture as of “primordial soup” for the developing of digital culture is analyzed in F. Turner. *From Counterculture to Cyberculture, Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, University of Chicago Press, Chicago, 2008. Stewart Brand, the founder of fundamental for both cultures *Whole Earth Catalog*, in a seminal article *Spacewar* on *Rolling Stone* drew a direct inheritance between psychedelics and information technologies: «Ready or not, computers are coming to the people. That’s good news, maybe the best since psychedelics». Brand describes also the hackers as direct heir of the hippies. S. Brand, *Spacewar*, in «*Rolling Stones*», 07/12/1972.

to find seeds of this crisis in the heart of the concrete development of the category of the Self and his features. The critique has profound consequences for the concepts of the moral agency that affects the political sphere. Who are the citizens, and who is accountable for his actions in our globalized and technological democracies? A particularly useful example of in these matters, alongside another study on the phenomenon of possession, better analyzed below, are the research carried out in the book *Vampyr. Storia naturale della resurrezione*⁶ by F.P. De Ceglia, a study focused on the origins of the belief in vampirism, tracing its genealogy up to the events related to the dissemination of the news by newspapers of the time that, on Christmas of 1731, the dead had risen and decided “to wage war” on the living.

The analysis of this episode of “natural history” of the European Modernity revealed an intertwining of phenomena that are far from being anthropologically “neutral” or simply relegable to folklore. At the heart of the belief in vampires there are religious-cultural⁵ and political-economic issues; ways of sociality modalities; climatic conditions (the long periods of frost characterizing the seasonal cycles in the countries where the vampirism epidemic originated) and ancestral worldviews and explanations of natural phenomena (the *vukodlaci* wolves chasing clouds in the sky)⁷. These ideas were widespread not only among the rural or poorer strata of the population but they also found fertile ground for their dissemination in the philosophical, theological, and scientific culture of the time and had deep political effects.

The analysis of vampirism – and the blurring of conceptual dichotomies taken for granted, such as life/death, human/non-human, mechanistic/animated, natural/unnatural, etc. – shows how the categories of modern philosophy, *in primis* the Self, need to be radically questioned. There is a fracture in the Modern Subject: monolithic and individual, unique and identical to itself, understood as a Cartesian thinking substance, ontologically hyposatized and separated from the world – or as a transcendental “I”

⁶ F.P. De Ceglia, *Vampyr. Storia naturale della resurrezione*, Einaudi, Torino 2023.

⁷ Cfr. Ivi, Ch. 5.

that gazes upon the world as a «judge of the real and the valid»⁸. This fracture opens up a conceptual space for a fruitful contamination of methods and objects between philosophical reflection on *Information*⁹ and fields of inquiry that might initially seem alien to it, but only in the case of an exclusively “engineering” approach to ICT technologies that neglects their intimate sociotechnical nature. Historical and socio-anthropological disciplines are essential to develop the ontological and epistemological rethinking that the digital revolution entails, as will be briefly illustrated in some thematic points throughout this work.

In the horizon of the digital revolution, an additional step is taken towards the crisis of modern subjectivity. This crisis is one of the keys to understanding the philosophical, socio-anthropological, ethical, and political consequences of the transformation due to ICT in their full “revolutionary” scope. The impact changes the status and response to questions such as: “What is the *τί ἐστί* of Man?”; “Who or what holds rights and duties, who or what is a moral agent or patient?”; “Who or what forms a social collective?” After the Fourth Revolution, it is no longer possible to pose these questions, nor to answer them, in the same way as the Moderns did.

2. Algorithms as Cultural Machines

Is necessary, at the macro-level of analysis, considerate the never neutral nature of the algorithm. Algorithms are cultural machines¹⁰. They incorporate, convey, multiply, and implement structures and practices of power, desire and reproduction of economic and political interests. They are based on and feed off the *symbol-language-magical thought* nexus in an interactive relationship between deep anthropological in-

⁸ «Il soggetto pensante che giudica ciò che è reale e valido» (P. Pecere, *Il dio che danza. Viaggi, trance, trasformazioni*, Nottetempo, Milano, 2021, p. 73).

⁹ Here the term “Information” is used as an umbrella term to indicate cyber-philosophy, digital philosophy, computational philosophy, AI philosophy and so on.

¹⁰ E. Finn, *What Algorithms Want: Imagination in the Age of Computing*, MIT Press, Cambridge Massachusetts 2017, p. 26.

stances and our cognitive structures¹¹. «From its bones in set theory and symbolic logic to the latest articulations of data-driven machine learning, computation casts a cultural shadow that is informed by this long tradition of magical thinking»¹². The digital is a *weltanschauung*¹³, as its theoretical antecedent is: the computational thought. It is no coincidence that the “grandfather” of AI was Thomas Hobbes¹⁴, who in the *Leviathan* proposed the first definition of thought as calculation¹⁵. Hobbes bases his entire political and scientific theory on

¹¹ A fascinating analysis of this relationship is found in the 1992 cyberpunk cult novel *Snow Crash* by Neal Stephenson, the first novel in which the word “Metaverse” appears. See, Finn, *What Algorithms Want*, cit., Chapter I. The relationship between symbol, syntax, and reality is also the theme of the long-standing problem of the Philosophy of Information (L. Floridi, «*Is semantic information meaningful data?*», «Philosophy and Phenomenological Research», 70, 2005, pp. 351-370), the *Symbol Grounding Problem* (S. Harnad, *The Symbol Grounding Problem*, «Physica», 42, 1990, pp. 335-346), which concerns how symbols used by a cognitive system can acquire meaning connected to the external world without necessarily «parasitizing» it from our mind.

¹² E. Finn, *What Algorithms Want*, cit., p. 2.

¹³ For a passionate critique of the world view conveyed by the digital, in line with the seminal work of Weizenbaum, (J. Weizenbaum, *Computer Power and Human Reason: From Judgment To Calculation*. Freeman, San Francisco 1976), more significant because it comes from a pioneer of the computer revolution, see J. Lanier, *You are not a gadget: A Manifesto*, A. Knopf, New York 2010.

¹⁴ Cfr. J. Haugeland, *Artificial Intelligence: The Very Idea*, MIT Press, Cambridge-Massachusetts 1985, p. 23.

¹⁵ «When a man reasoneth, he does nothing else but conceive a sum total, from addition of parcels; or conceive a remainder, from subtraction of one sum from another: which (if it be done by words,) is conceiving of the consequence from the names of all the parts, to the name of the whole; or from the names of the whole and one part, to the name of the other part». T. Hobbes, *Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiastical and Civil Edited with an Introduction and Notes by C. A. Gaskin*, Oxford University Press, Oxford 1996, Book I, chap. V, p. 18 orig. The English philosopher one of the two distinguished thinkers – together with Boyle – who gives the name to the fundamental work of Shapin and Schaffer, *The Leviathan and the Air Pump*, is constructed, which thematizes and investigates the never neutral nature of the scientific enterprise. Knowledge and political dimension, for the two authors, are never separate or independent issues. «Knowledge, as much as the state, is the product of human actions. Hobbes was right». S. Shapin, S. Schaffer, *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*, Princeton University Press, Princeton NJ 1985, p. 344.

mathematical demonstration. In the *par excellence* work about Power, the computational power is a form of power. Hobbes structures his entire political and scientific theory on mathematical demonstration. Anthropological reflection on science can find here fertile ground to thematise the political nature of computation. Bruno Latour's words are illuminating:

He arrives at all his scientific results not by opinion, observation or revelation but by a mathematical demonstration, the only method of argument capable of compelling everyone's assent; and he accomplishes this demonstration not by making transcendental calculations, like Plato's King, but by using a purely computational instrument, the Mechanical Brain, a computer before its time. Even the famous social contract is only the sum of a calculation reached abruptly and simultaneously by all the terrorized citizens who are seeking to liberate themselves from the state of nature¹⁶.

Without delving into the ontological interpretations of *Hypercomputation* (see *supra* in note), the algorithmic thought, from what G. Dyson has named *the Turing's Cathedral*¹⁷, sanctions and conveys the computationalist vision: «Algorithmic thinking encodes the computationalist vision, the maximalist idea that all complex systems will eventually be made equivalent through computational representation»¹⁸. The power of cut and paste in the digital realm allows us to cut and re-paste our reality and our ideas about Modernity¹⁹ creating, recreating, or revealing entities made of both nature and culture denied by it. In this process, in the Fourth Revolution, the fiction of a clear separation and opposition – and in this fiction lies the fundamental trait of the modern age – between the world and man, between nature and culture, collapses, revealing the presence of *hybrids* and allowing their conceptualisation.

¹⁶ B. Latour, *We have never been modern*, Harvard University Press, Cambridge 1993, p. 19.

¹⁷ G. Dyson, *Turing's Cathedral: The Origins of the Digital Universe*, Vintage Books, New York 2012, chapter 13.

¹⁸ E. Finn, *What Algorithms Want*, cit., p. 43.

¹⁹ L. Floridi, *The Ethics of Artificial Intelligence. Principles, Challenges and Opportunities*, Oxford University Press, Oxford 2023, p. 4.

3. There Is a Revolution. Who Is The Revolutionary?

The very concept of the Fourth Revolution, structured and nourished by the computational power outlined above, challenges the idea of subjectivity inherited from modernity. The revolution led by Alan Turing – following Copernicus, Darwin, and Freud, symbols of previous revolutions²⁰ - marks another step in the process of rethinking by humans about the world and themselves. Gradually, Man is no longer at the centre of Creation (*1st Revolution*); he is no longer the pinnacle, unnaturally separate, of the living nature (*2nd Revolution*); he is no longer Cartesian clear, univocal, and self-transparent thought (*3rd Revolution*). Turing's man, finally, must renounce the primacy/monopoly on information processing²¹ and intelligent action. He must

²⁰ Cfr L. Floridi, *The Fourth Revolution. How the Infosphere is Reshaping Human Reality*, Oxford University Press, Oxford 2014, pp. 87-94.

²¹ For information here does not mean Shannon's Information. Shannon's information is quantitative, free from the semantic dimension that it can assume, understood in relation to the measurement of the entropy and uncertainty of a message and it can be expressed with the average logarithm of the message's improbability; formalized as: $H = - \sum p_j \log_2 p_j$ (H = amount of information, *the entropy*; Σ = the sum over all messages j ; p_j = probability of message j) a formula that represents a decisive step forward compared to Hartley's "classic" $H = n \log s$ (s = number of possible symbols, n = number of transmitted symbols). (see R. Hartley, *Transmission of informations*, «Bell System Technical Journal» (July) 7 1928, pp. 535-563; C.E. Shannon, *A Mathematical Theory of Communication*, «Bell System Technical Journal», Vol. 27, 1948. pp. 379-423, 623-656. An effective and non-technical introduction can be found in J. Gleick, *The Information: A History, a Theory, a Flood*, Pantheon Books, New York 2001, for an overview of the concept of information see P. Adriaans, P. Information, in *Stanford Encyclopedia of Philosophy*, 2012/2020. The information taken into account here is semantic information, in its "strong" sense. Considered: SDI (standard definition of information); DOS (declarative, objective, semantic); σ (*infor*) = symbol used to refer to discrete elements of information, regardless of their semiotic code and physical implementation; δ = data not alethically qualified, well formed and endowed with meaning: Strong semantic information in the sense (RSDI) will thus be defined: RSDI σ is an instance of DOS information if and only if: 1. σ consists of n data (d), for $n \geq 1$; 2. the data are well formed (wfd); 3. the wfd are meaningful ($mwfd = \delta$); 4. the δ are truthful. (L. Floridi, *The Philosophy of Information*, cit., pp. 83-83; criticized in M.T. Ferguson, *Two paradoxes of semantic information*, «Synthese» 192(11), pp. 3719-3730. and in B. Lundgren, *Does semantic information need to be truthful?*, «Synthese» Vol. 196, No. 7, Special Issue on Between First- and Second-Order Logic July 2019, pp. 2885-2906.

share the world and the infosphere with technologies, the ICTs, capable of supporting or even replacing him in complex²¹ activities. Activities that have historically been considered exclusively human prerogatives. *Homo Sapiens*, through ICT, also discovers himself as an “*Inforg*”, an informational organism, living “*Onlife*” (without a discontinuity between online and offline) in a “*hyper-historical*”²² society – socially and individually dependent largely on ICT – where digital technologies become technologies of the Self, capable of constructing, shaping, transforming, and reconfiguring our personal identity²³.

The gap from the ontological hypostatisation of the Cartesian *res cogitans* or the Kantian transcendental *a priori* could not be more profound. ICT highlight the nature of the Self as a Multi-Agent System (MAS), constituted by different dimensions/membranes (bodily, cognitive, and conscious) interacting and influencing each other in a poetic process²⁴. In the *Phaedrus*, Plato had already characterized the nature of the Self as at least tripartite with the famous metaphor of the charioteer and the two horses²⁵. Digital technologies solve the problems posed by recognizing the non-substantial, non-univocal, and non- diachronically stable nature of the Self, as pointed out by empiricist philosophers like John Locke and David Hume in their cri-

²² L. Floridi, *The Fourth Revolution*, cit., p. 94.

²³ Cfr. Ivi, Chapter I, pp. 59-68, 94-96.

²⁴ Id., *The Informational Nature of Personal Identity*, «Minds and Machines», 21, 4/2011, pp. 549-566.

²⁵ «In my analogy, a soul is like an organic whole made up of a charioteer and his team of horses. Now, while the horses and charioteers of gods are always thoroughly good, those of everyone else are but a mixture. Although our inner ruler drives a pair of horses, only one of his horses is thoroughly noble and good, while the other is thoroughly the opposite. This inevitably makes driving, in our case, difficult and disagreeable» (Plato, *Phaedrus*, 246a-b; translated with an Introduction and Notes by R. Waterfield, Oxford University Press, Oxford 2003, p. 28. From an informational standpoint, the Platonic analogy could be seen as a technological analysis of a multi-agent system (MAS), where the issue concerns the relationship among the three parts. This poses questions about continuity, coordination, and the unity of the MAS. This might suggest that even the chariot itself, in its functional parts, could be considered as the fourth element of the Platonic self-system, as it is what holds together and coordinates the various subunits. L. Floridi, *The Informational Nature of Personal Identity*, cit.

tique of the subject as *res cogitans*. The Self that emerges from informational analysis is neither a substance separated from the world nor a mere *a priori* epistemic function guaranteeing the objectivity of sensations. The subject, and its identity, far from being always unique and identical to themselves, are constructed and produced in relation to the world and other selves, each as designers of narratives continually externalized and related to other selves, in an interactive relationship that ICT amplifies and deepens enormously.

4. Intelligence: Imitation, Behaviour, Function

The birth act of AI – the famous *Proposal for the Dartmouth Summer Research Project on Artificial Intelligence* organized by John McCarthy, with the participation of Minsky, Rochester, Shannon, Simon, Newell et al. – establishes Artificial Intelligence as a counterfactual: «For the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving»²⁶. The first “cry” of AI – or rather the first in which the name *Artificial Intelligence* appears – does not define Artificial Intelligence, nor does it indicate how to recreate it, in whole or in part, at a human level, but rather defines a behaviour²⁷.

²⁶ J. McCarthy, M. Minsky, et al., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, «AI magazine», n. 27, August 31, 1955. For the analysis of AI as counterfactual, L. Floridi, *The Ethics of Artificial Intelligence*, cit., pp. 16-20.

²⁷ The history of AI as a science is structured around this problem. Both in terms of its development as a science - the succession of Winters and Summers (increase or decrease in interest, public and private funding, public relevance of AI) demonstrating the fallacy of an internalist approach to this discipline - and from the perspective of research paradigms (strong/weak AI; engineering/production AI; symbolic/subsymbolic AI) and also at the level of theoretical approach, based on Thought or Behavior, Rationalist/Human-centric. Cfr. M. Mitchell, *Artificial Intelligence: A Guide for Thinking Humans* (First ed.). Farrar, Straus and Giroux, New York 2019, Part I; S. Russell, P. Norvig, *Artificial Intelligence. A Modern Approach* (Fourth Edition), Pearson, London 2021, Ch. 1.1.

Intelligence is behaviour. The digital subject, is a multi-agent system of interactive behaviours with each other and with the environment (the world, or rather the *Infosphere*) in a co-constitutive relationship of multiple identities, themselves subject to the technological shaping power of ICT. Without delving further into the subject, the counterfactual nature of the AI definition by its “godfathers” already marks a profound departure from the traditional conception of agency, which was traditionally ascribed only to the human subject. Machines could also act and could do so like humans. The Dartmouth definition pairs with the definition of the *Imitation Game*, the so-called *Turing Test*²⁸, proposed by Alan Turing in the 1950 in his classic paper *Computing Machinery and Intelligence*²⁹. The English mathematician states regarding the question: «Can machines think?», «I believe this question to be too meaningless to deserve discussion». Using a popular role-playing game

²⁸ Unfortunately, it is not possible here to critically discuss Turing’s famous proposal and the vast literature on it. For a general overview – from the 1950 proposal of the Imitation Game to the recent editions of the Loebner Prize at the University of Reading (an annual chatbot contest competing to achieve the highest score in the Turing Test) – and a discussion of criticisms (such as the famous ones from Searle’s Chinese room or Ned Block’s blockhead), see G. Oppy, D. Dowe, *The Turing Test*, in *Stanford Encyclopedia of Philosophy*, 2003/2021. An effective philosophical examination is found in J.H. Moor, *An Analysis Of the Turing Test*, «Philosophical Studies», 30, 1976, pp. 249-257 and J.H. Moor, *The Status and Future of the Turing Test*, «Minds and Machines», 11, 2001, pp. 77-93. A detailed critique of the inefficacy of the Turing test as an effective test to discriminate the presumed intelligence of machines, and the need to replace it with others (related to semantic dimension, understanding of contexts and common sense, or the ability of detachment) more effective given the technological development of new chatbots and LLMs is found in Larsson 2021. E.J. Larsson, *The Myth of Artificial Intelligence: Why Computers Can’t Think the Way We Do*, Harvard University Press, Cambridge 2021, Chapter 5, 13. It’s mandatory to mention the *Voight-Kampff Test*, based on empathic ability, developed by Philip K. Dick in his 1968 science fiction work *Do Androids Dream of Electric Sheep?* which inspired Ridley Scott’s 1982 film *Blade Runner*. Moreover, each of us, probably several times a day, undergoes a Turing Test, in which, paradoxically, the machine judges our humanity, when it encounters the now familiar screen containing a *CAPTCHA* (Completely Automated Public Turing test to tell Computers and Humans Apart). To date, *CAPTCHA* is an insurmountable task for machines, even for the most advanced AI models.

²⁹ A.M. Turing, *Computing Machinery and Intelligence*, «Mind», 59, 236, Oct. 1950, pp. 433-460.

of the time – involving three participants, A (Man), B (Woman), and C (Judge), who is separated from the other two and must determine who is the man and who is the woman through simple written questions and answers – Turing reformulates the “meaningless” question about machine thinking into a verifiable scientific hypothesis: «Can a machine participate in the Imitation Game?»³⁰. If the machine deceives C, it should be attributed intelligence.

The boundaries between machine and man blur and become a function of the machine's ability to produce an action that can be understood as intelligent by a human being. There is no longer anything of the typical characteristics of modern subjectivity (and, along with it, agency). The distance from Descartes could not be greater³¹. Strictly

³⁰ Ivi.

³¹ Ironically, the Turing test has been called a “Cartesian” test. In both the *Discourse on Method* and the *Meditations*, there are precedents of the Imitation Game with mentioning: «At this point I had dwelt on this issue to show that if there were such machines having the organs and outward shape of a monkey or any other irrational animal, we would have no means of knowing that they were not of exactly the same nature as these animals, whereas, if any such machines resembled us in body and imitated our actions insofar as this was practically possible, we should still have two very certain means of recognizing that they were not, for all that, real human beings. The first is that they would never be able to use words or other signs by composing them as we do to declare our thoughts to others. For we can well conceive of a machine made in such a way that it emits words, and even utters them about bodily actions which bring about some corresponding change in its organs (if, for example, we touch it on a given spot, it will ask what we want of it; or if we touch it somewhere else, it will cry out that we are hurting it, and so on); but it is not conceivable that it should put these words in different orders to correspond to the meaning of things said in its presence, as even the most dull-witted of men can do. And the second means is that, although such machines might do many things as well or even better than any of us, they would inevitably fail to do some others, by which we would discover that they did not act consciously, but only because their organs were disposed in a certain way. For, whereas reason is a universal instrument which can operate in all sorts of situations, their organs have to have a particular disposition for each particular action, from which it follows that it is practically impossible for there to be enough different organs in a machine to cause it to act in all of life's occurrences in the same way that our reason causes us to act», R. Descartes, *Discourse on Method*. Part V, in *A Discourse on Method*, translated by I. McLean, Oxford University Press, Oxford 2006, pp. 46-47. The second place is in the *Second Meditation*: «from this I would have immediately concluded that I therefore knew the wax by

speaking – and not only retrospectively, since *General Artificial Intelligence* (AGI) is a far-off goal today, despite machines performing various activities that, from a certain perspective, can be considered intelligent or even creative (consider *Large Language Models* (LLM) like ChatGPT or image-generating AIs³² like Midjourney or DALL-E) – following Turing, to produce an intelligent action, thought is no longer necessary. Nor is any specific physical or organic materiality. Turing and his machines are a powerful vector for the idea that the mind is software and the brain nothing more than one of its possible hardwares. This idea developed into the philosophy of mind school of thought known as *Functionalism*, the theory according to which mental states are computational states, expressible in terms of formal rules for the manipulation of symbols. They express nothing more than functions. Therefore, they can be implemented by any system, including a digital machine, if that has the right characteristics.

5. Subject/Subjects: a Society, a Network

It is useful to point out – considered the bidirectional relationship explanandum/explanans between Artificial Intelligence and Models of

the sight of my eyes, not by the inspection of the mind alone-if I had not happened to glance out of the window at people walking along the street. Using the customary expression, I say that I “see” them, just as I ‘see’ the wax. But what do I actually see other than hats and coats, which could be covering automata? But I judge that they are people. And therefore what I thought I saw with my eyes, I in fact grasp only by the faculty of judging that is in my mind», Descartes, *Meditations on First Philosophy, Second Meditation*, p. 32, in *Meditations on First Philosophy With Selections from the Objections and Replies*, Translated with an Introduction and Notes by M. Moriarty, Oxford University Press, Oxford 2008, p. 23. It should be noted that in *Calculating Machines and Intelligence*, Turing addressing possible objections to the Imitation Game, tackles philosophical topics of great philosophical importance such as: the dualism of substance; solipsism; biological/cultural determinism; self-awareness. Cfr. A.M. Turing, *Computing Machinery and Intelligence*, cit..

³² Briefly, A Large Language Model (LLM) is an artificial intelligence – or, better term a computational statistical algorithm self-supervised or semi-supervised – system trained on extensive text datasets to understand and generate human-like language. Generative AI refers to artificial intelligence systems designed to create new content, such as text, images, or music, by learning patterns from existing data.

the Mind – that within the computationalist thought itself arise a fracture of the subject's unity. Marvin Minsky, a pioneer of AI, explicitly speaks of *Society of Mind*³³. Once that is established equivalence “mind = brain”, for Minsky there would be a myriad of agents that are mindless, but from whose ‘social’ interaction the mind would result. The relationship between these agents parts of the mind «is sometimes one of cooperation, but more often of conflict»³⁴. The consequence is that human intelligence, consciousness, and mind are not reducible to a single and simple principle. On the contrary, far from there being «magical tricks»³⁵. The power of intelligence derives from our vast diversity, not from a single perfect principle. At the very heart of consciousness, understood computationally, there is a decisive feature of fragmentation.

Similarly, considering the mind as the ability to manipulate symbols through rules – a fundamental definition of the symbolic approach to artificial intelligence, as Haugeland does concerning GO-FAI (*Good Old-Fashioned Artificial Intelligence*)³⁶ – one encounters the «paradox of mechanical reason»³⁷. The presence of an intelligent, not merely mechanical, *homunculus* manipulating symbols necessarily implies, and here lies the nature of the paradox, an infinite regressive multiplication of *homunculi* manipulating symbols.

It is beyond the scope of this paper to further analyse this paradox. But it should be noted how the computationalist theory itself infinitely multiplies the population within the mind. The difference with the modern consideration of subjectivity, identity, and consciousness

³³ M. Minsky, *The Society of Mind*, Simon & Schuster, New York 1985.

³⁴ «Il cervello umano è una vasta società organizzata, composta di molte parti diverse. Dentro il cranio dell'uomo sono stipati centinaia di tipi diversi di motori e organizzazioni, meravigliosi sistemi evolutisi e accumulatisi nel corso di centinaia di milioni di anni. Alcuni di questi sistemi, ad esempio le parti del cervello che ci fanno respirare, funzionano in modo pressoché indipendente. Ma nella maggioranza dei casi queste parti di mente devono convivere con le altre, in un rapporto che è a volte di collaborazione, ma più spesso di conflitto». M. Minsky, *Prefazione all'edizione italiana*, in *La società della mente*, Adelphi, Milano 1989, p. 20.

³⁵ M. Minsky, *The Society of Mind*, cit., p. 308. «What magical trick makes us intelligent? The trick is that there is no trick».

³⁶ J. Haugeland, *The Artificial Intelligence*, cit., p. 112.

³⁷ Cfr. Ivi, pp. 36-44.

sharpens further if one briefly considers the subsymbolic paradigm of AI and its corresponding theory of mind, *Connectionism*. These are founded on the dimension of mental operations (e.g., perception) that lie below the level of logical-symbolic functions (e.g., the semantics of language). Considering the cardinal concept of this paradigm, *Artificial Neural Networks* (ANN) – an explanatory model of the mind and a key to the astounding success of Machine Learning (ML) and Deep Learning (DL)³⁸ – despite the recognizable importance of inductive

³⁸ Unfortunately, it is not possible to delve further into the concepts underlying the Machine Learning revolution here. ML is founded on the insights of Rosenblatt’s Perceptron (F. Rosenblatt, *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*, «Cornell Aeronautical Laboratory, Psychological Review», v65, No. 6 1958, pp. 386–408), refuted by Minsky and Papert (M. Minsky, S. Papert, *Perceptrons*, MIT Press, Cambridge, MA 1969) an event that stalled research on neural networks for many years. Key milestones in the resurgence of this research program (whose success is evidenced by the fact that in public discourse now AI and ML/DL are used synonymously) have been research on *Parallel Distributed Processing Neural Networks* (PDP) (J.L. McClelland, D.E. Rumelhardt PDP Research Group (1986) *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, 2, 216-271) (which in turn opened the way to the study of parallel functioning of human brain structures) and the spectacular victory of the ML algorithm *AlphaGo* in 2016 against the Go champion Lee Sedol (an event that surpasses in historical-technological importance even the victory of *Deep Blue* against Kasparov in 1997). On Connectionism as a branch of philosophy of mind that uses Artificial Neural Networks as an explanatory model, see C. Stinson, *Explanation and connectionist models* and C. Buckner, & J. Garson J. (2018), *Connectionism and post-connectionist models* in M. Spreveak, M. Colombo, *The Routledge Handbook Of The Computational Mind*, Routledge, Londra, pp. 120-133, 76-90; P. Smolensky, *On the proper treatment of connectionism*, in «Behavioral and Brain Sciences», vol. 11, Cambridge University Press, Cambridge 1989. For an overview of ML and DL, see M. Domingos, *The Master Algorithm. How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, New York 2015 (although inspired by an almost “messianic” confidence in the potential of machine learning); M. Mitchell, *Artificial Intelligence*, cit., Ch. II and VI. Kai-Fu Lee and Chen Quifan offer a suggestive explanation of DL alongside a realistic science fiction narrative. Kai-Fu Lee, Chen Quifan *AI 2041. Ten Visions for Our Future*, Penguin, New York 2201, Ch. I. The treatment in S. Russell, P. Norvig, (2021), *Artificial Intelligence*, cit., Ch. V is more technical. An effective critique of the issues with ML/DL is in G. Marcus, *Deep Learning: A Critical Appraisal*, arXiv, January 2, 2018. On algorithmic black boxes, also from a socio-political and legal standpoint, see F. Pasquale, *The Black Box Society The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge Massachusetts 2015.

inference, the processes leading to intelligent output or the identification of patterns from input data are by definition unknowable. They occur in the black box of nodes' convolution at different levels in the hidden layers of neural networks.

The astonishing successes of Machine Learning strike a further blow to one of the major claims of the modern subject: its supposed self-transparency and simplicity.

6. The "I": a Breathe?

The impact of digital technology and the recognition of the informational nature of human identity revitalize contemporary discussions around certain topics that crystallise ancient hopes or ancestral fears³⁹

³⁹ Contemporary fears about potential *Existential Risks* (X-Risks) to the human species due to superintelligent AI (the one described in N. Bostrom, *Superintelligence. Paths, Dangers, Strategies*, Oxford University Press, Oxford 2014) - exemplified by the now famous *Pause Giant AI Experiments: An Open Letter* from March 2023 by the *Future of Life Institute* (and signed for various reasons by many representatives, researchers, academics, and entrepreneurs from the AI world) - converge with what is the latest philosophy *à la mode* in Silicon Valley: *Longtermism*. Based on *Effective Altruism* (using scientific evidence and reason to understand and act on how to maximize benefits for others, maximizing the quantity of good, e.g., through rigorous cost-benefit scientific analysis), Longtermism asserts that the most important determinant of the value of our actions today is how these actions – which must quantitatively maximize. Positive consequences for human beings – influence the long/very long-term future. Helping wealthy nations, which know how to better use resources, rather than poor ones, is more important from a longtermist perspective. Avoiding the X-Risk of a super-AI in a remote future is more important than fighting climate change or AIDS today. Consuming Earth's resources so that the galaxy is colonized by trillions of human beings for longtermists is an ethically sustainable option. The main texts that underpin the longtermist view, which is gathered in think tanks like the *Future of Life Institute* or in universities like Oxford - and which also holds important advisory positions in the highest global political bodies - are W. MacAskill, *What we owe the Future*, Basic Books, New York 2022; T. Ord, *The Precipice: Existential Risk and the Future of Humanity*, Bloomsbury, Londra 2020; M Tegmark, *Life 3.0. Being Human in the Age of Artificial Intelligence* Knopf, New York 2017. An effective critique, which also collects the most disconcerting statements of longtermists, is E. P. Torres, *Against Longtermism*, in «Aeon», 19 October 2021. Curiously, these futuristic theories bear a resemblance to one of the more "modern" traditions of philosophical

(these topics can be examined, cause their very nature, by the anthropological disciplines). These include the possibility of enhancing the physical/intellectual capabilities of human beings (e.g., *Transhumanism*⁴⁰) or even achieving immortality (through the upload of consciousness into natural or artificial physical supports different from the original human body)⁴¹.

These theoretical perspectives once again demonstrate how the Subject, an heir to philosophical modernity, is being radically questioned. They highlight, by now unsurprising, a “family resemblance” between the points examined and instances of conceptualizations of the Subject, as claimed above, that we can find in instances of anthropological and historical studies. If the analysis of vampirism⁴² shows the crisis of modern subjectivity, the same does the one about the theme of *possession*.

Both these studies deal with anthropological entities vastly different from those postulated by doctrines of the “Modern” Man. They reveal the inherently “unnatural” character of this concept and its essential peculiarities: the indivisible substantial or functional unity of the individual and its role as a structuring instance of the world. «It is indeed appropriate to remember how ‘unnatural’ and largely confined to the centuries of the ancient régime is the idea of a monolithic individual consciousness: a narrative centre of gravity»⁴³.

thought, good old utilitarianism. From this genealogy, it is easy to understand the economic interests lying behind Longtermism.

⁴⁰ A sharp and detailed critical overview of the fallacies of transhumanism is the provided in L. Dibattista, *La filosofia di Altered Carbon, ovvero: le fallacie del Transumanesimo*, «Idee in form@zione», 2024, pp. 19-35.

⁴¹ It is interesting to note how, apart from the plausibility of these options in a radical functionalist perspective, the descent of these theses from the “classic” considerations of criticism of the Cartesian substantial subject such as those contained in Locke’s *Essay on Human Understanding*, in his famous thought experiments about personal identity, consciousness, qualitative identity, and memory.

⁴² Is remarkable the impact that journalistic reconstructions have had on this matter, as an example of the power of information technologies even in “analog” form, such as a daily newspaper or magazine.

⁴³ F. P. de Ceglia, *Vampyr*, cit., p. 106. «È in effetti opportuno ricordare quanto «innaturale» e in linea di massima circoscritta ai secoli di antico regime sia l’idea di una monolitica coscienza individuale: un centro di gravità narrativo».

The modern subject is an intellectual construct that conveys, in a certain prescriptive description of the human, a very specific vision of the world, society, and power relations. It is a «perspective [...] counterintuitive. And far from being widely shared outside the homilies of priests or the treatises of certain philosophers»⁴⁴. The belief in the plausibility of vampirism was nurtured by widespread conceptions about the nature of the self that are almost the same as those emerging from its informational consideration, such as the fragmented, non-unique nature of the Self, which instead reveals itself as a MAS. «The belief that some living vampire-like being possessed a double soul – a sort of double personality, to modernize, or a dual way of relating to the world – was not so far-fetched»⁴⁵. Moreover, this belief was not only widespread among the superstitious populace but also among «men of culture, even in the 'enlightened' Western Europe, who, following the Platonic example, spoke of the tripartition of souls»⁴⁶. The same is true for those who, advocating the most "orthodox" version of this belief in the multiple nature of the Self, had spoke of «functional articulations of the same physiological-behavioural spiritual principle, in the manner of Aristotle»⁴⁷. Beliefs in the plurality of souls are based, in addition to a «resemantization of Christianity»⁴⁸ on the intuitive realization that the self is never always the same. «I want to do one thing, but I do another; sometimes I am fully lucid, other times not; I act one way now, another way later; I ignore how my body functions, yet every organ knows what to do as if there were a little 'instructor' inside me, etc. All this could indeed give me the impression of hosting multiple spirits»⁴⁹.

⁴⁴ «Una prospettiva [...] poco intuitiva. E tantomeno condivisa al di fuori delle omelie dei sacerdoti o dei trattati di certi filosofi» (Ibidem).

⁴⁵ «La convinzione che qualche vampiroide vivo custodisse una doppia anima – una sorta di doppia personalità, per voler attualizzare, o un duplice modo di rapportarsi al mondo – non era quindi così balzana» (Ivi, p. 107).

⁴⁶ «Uomini di cultura, anche nella "illuminata" Europa occidentale, seguendo l'esempio platonico, parlavano di tripartizione delle anime» (Ibidem).

⁴⁷ «Articolazioni funzionali di un medesimo principio spiritual fisiologico-comportamentale, alla maniera di Aristotele» (Ibidem).

⁴⁸ «Risemantizzazione del cristianesimo» (Ibidem).

⁴⁹ Ibidem. «Io voglio fare una cosa, ma ne faccio un'altra; a volte sono pienamente lucido, altre no; agisco ora in un modo, ora in uno diverso; ignoro il funzionamento del mio corpo, eppure ogni organo sa quel che deve fare come se dentro di me ci

The Self is configured, not so much differently from what Sigmund Freud, the champion of the *Third Revolution*, would have argued, as an «ecosystem of forces»⁵⁰. This idea also has a basis in the “scientific” culture of the time. It «was especially not very far from what even certain modern Western medico-philosophical culture, with different words, would have been willing to admit»⁵¹. What has just been said about vampires aligns with the analysis of the phenomenon, which articulates the ethnographic-anthropological reflection within the historical-philosophical prospective, of possession.

Without delving into the specifics of the possessive phenomenon, we can also notice here the illusory nature of the claim that the subject is «an inviolable entity created by God, an immortal soul, the thinking subject who judges what is real and valid»⁵². Consider our knowledge of our own Self. It is neither the first nor the most evident: «Self-knowledge is not immediate and intuitive. Who indeed am I?»⁵³. The concept of Self resulting from the anthropological analysis, similarly to the one of the analytic philosophy⁵⁴, is «the subject of predicates. Only in the course of experience are these predicates gathered, and then one attempts to determine who or what I am individually, my identity»⁵⁵.

At any given moment, like any discrete state system, *prima facie* there is only the body, and just in relation with the body we can, have instances such as a self, a consciousness that has sensations, knowl-

fosse un piccolo «istruttore» ecc. Tutto ciò potrebbe insomma darmi l'impressione di ospitare più spiriti».

⁵⁰ «Ecosistema di forze» (Ibidem).

⁵¹ «Soprattutto non era molto lontana da ciò che anche certa cultura medico-filosofica occidentale di età moderna, con parole differenti, sarebbe stata ben disposta ad ammettere» (Ibidem).

⁵² «Un'entità inviolabile creata da Dio, un'anima immortale, il soggetto pensante che giudica ciò che è reale e valido» (P. Pecere, *Il dio che danza*, cit., p. 73).

⁵³ «La conoscenza di noi stessi non è immediata e intuitiva. Chi sono infatti io?» (P. Pecere, *Il dio che danza*, cit., p. 208).

⁵⁴ Cfr. M. Di Francesco, *Introduzione a Russell*, Laterza, Bari 1990, p. 98.

⁵⁵ «Io è un pronome che designa “una cosa dal significato indeterminato”, cioè il soggetto dei predicati. Solo nel corso dell'esperienza si raccolgono questi predicati e si tenta quindi di determinare chi o cosa sia io individualmente, la mia identità» (P. Pecere, *Il dio che danza*, cit., p. 208).

edge, memories. But all the predicates and entities constituting identity never have a fixed and established character forever. The Self tries to build a narrative description «of its own person using its body and language to connect the elements of lived experience into a story, but the body changes, and so do the content and meaning of the story narrated by the self»⁵⁶.

This prospective cannot be any closer to the poietic-narrative concept resulting from the informational analyses of the Self⁵⁷. Thus,

⁵⁶ «Il sé cerca di costruire una descrizione della propria persona usando il proprio corpo e il linguaggio per collegare gli elementi del vissuto in una storia, ma il corpo cambia, e pure il contenuto e il senso della storia narrata dal sé mutano continuamente, per gradi o per svolte radicali» (Ibidem).

⁵⁷ Cfr. *supra* respect the constitution of the Self by Information philosophy: «In order to emerge and flourish, the mind needs to make sense of its environment by continuously investing [...] with meaning. Mental life is thus the result of a successful reaction to a primary *horror vacui semantici*: meaningless (in the non-existentialist sense of 'not-yet-meaningful') chaos threatens to tear the Self asunder, to drown it in an alienating otherness perceived by the Self as nothingness, and this primordial dread of annihilation urges the Self to go on filling any semantically empty space with whatever meaning the Self can muster, as successfully as the cluster of contextual constraints, affordances, and the development of culture permit. This giving meaning to, and making sense of reality (semanticization of Being), or reaction of the Self to the non-Self (to phrase it in Fichtean terms), consists in the inheritance and further elaboration, maintenance, and refinement of factual narratives: personal identity, ordinary experience, communityethos, family values, scientific theories, common-sense-constituting beliefs, and so forth». L. Floridi, *The Philosophy of Information*, cit., p. 7. Informationally, the Self, personal identity, can be considered akin to Proust's narrative artifacts: «Then there is a second approach, more recent, known as the Narrative theory of the self. According to it, your identity is a 'story', understood as a socio- and/or auto-biographical artefact. Recall what Proust said about the social self. We 'identify' (provide identities to) each other, and this is a crucial, although not the only, variable in the complex game of the construction of personal identities, especially when the opportunities to socialize are multiplied and modified by new ICTs». L. Floridi, *The Fourth Revolution*, cit., pp. 68-69, translated to tackle the Self in this prospective, in a continuous co-constructive process (amplified by ICTs), the appropriate LoA is necessary. «Identity and sameness relations are satisfied according to the LoAs adopted, and these, in turn, depend on the goals being pursued. This is not relativism: given a particular goal, one LoA is better than another, and questions will receive better or worse answers. The ship will be Theseus's, no matter how many bits one replaces, if the question is about legal ownership (try a Theseus trick with the taxman); it is already a different ship, for which the

it becomes evident how the subject is not a substantial, naturalised, or timeless entity, but a historical product, the result of a particular culture, influenced, in turn, by other worldviews.

The psyche was initially a vital ‘breath’, the one that is extinguished with the last breath (an etymology similar to the one some scholars find in the link between the Sanskrit *ātman* and the German *atmen*, ‘to breathe’); later, philosophy would distinguish the parts of the soul corresponding to various vital, emotional, and cognitive functions⁵⁸.

The consequences also involve another prerogative traditionally attributed to the Subject by the philosophical modernity: Responsibility. Even within the Western tradition, this claim is unfounded: «The Homeric heroes often said they acted under the influence of a god or a demon, and only gradually did a sense of full individual responsibility develop»⁵⁹. In conclusion, applying anthropological analysis to the Self, if «we consider globally the different perspectives of human civilisations, we realise that the person understood as a «bounded, unique, more or less integrated universe of motivation and cognition’ is an exception»⁶⁰.

7. Moral Agency in a Democracy of Things

We can briefly draw the ethical consequences affecting the definition of political society. The modern subject has long been considered the only intelligent agent and, because of his monopoly of intelligence

collector will not pay the same price, if all one cares about are the original planks». L. Floridi, *The Informational Nature of Personal Identity*, cit.

⁵⁸ «La *psyché* era inizialmente un “soffio” vitale, quello che si estingue con l’ultimo respiro (un’etimologia somigliante a quella che alcuni studiosi ritrovano nel nesso tra il sanscrito *ātman* e il tedesco *atmen*, “respirare”); in seguito la filosofia avrebbe distinto le parti dell’anima corrispondenti ad altrettante funzioni vitali, emotive e cognitive» (P. Pecere, *Il dio che danza*, cit., p. 73).

⁵⁹ «Gli eroi omerici dicono spesso di aver agito per l’intervento di un dio o di un demone, e solo gradualmente si sviluppa un senso di piena responsabilità individuale» (Ibidem).

⁶⁰ «Consideriamo globalmente le diverse prospettive delle civiltà umane ci accorgiamo che la persona intesa come “universo di motivazione e cognitivo delimitato, unico, più o meno integrato” è un’eccezione» (Ivi, p. 74).

and rationality, the only qualified moral agent. Only to him pertains to have agency, accountability, responsibility, and, Kantianly, freedom as a necessary condition of possibility for these categories of moral action. However, because the effect of the re-ontologizing and re-epistemologizing impact of ICT on the world and man, now the machines, at an appropriate Level of Abstraction (LoA) – constituted by *interactivity*, *adaptability*, and *autonomy*⁶¹ – must also be considered moral agents producing ethically relevant consequences. Machines must be regarded as qualified agents, capable to participate to the “moral game”⁶² in the world of the Fourth Revolution. The modern subject and its traditional concept of attribution of the moral agency – reducing morality to individual responsibility, attributing this only to the mental sphere, and making this sphere an exclusive human prerogative – have been an obstacle to the development of an informational ethical theory capable of overcoming the narrow confines of a “mi-

⁶¹ The three characteristics represent the minimal Level of Abstraction (LoA) required to consider artificial agents as moral agents. An LoA refers to a set of relevant properties (observable, i.e., typified and interpreted variants) of a system that are selected to describe, analyze, and understand that system in a given context. Cfr. L. Floridi, *Logic of Information*, Oxford University Press, Oxford 2019, pp. 41-47. In general, on the concept of the Level of Abstraction, see L. Floridi, *The Method of Levels of Abstraction*, in «Minds and Machines», 18, 3, 2008, pp. 303-329. Specifically, the three characteristics of: Interactivity (1), Autonomy (2), Adaptability (3) mean: (1) that the agent and its environment can act upon each other. Typical examples include the input or output of a value or the simultaneous execution of an action by the agent and the patient, such as gravitational force between bodies; (2) that the agent is able to change its state without directly responding to the interaction. Direct response to the interaction: it can execute internal transitions to change its state. Thus, an agent must have at least two states. This property grants the agent a certain degree of complexity and independence from the environment; (3) that the agent's interactions can modify the transition rules by which it changes state. This property ensures that an agent can be seen, at a given LoA, as an agent that learns to operate in a way that critically depends on its experience. It should be noted that if the transition rules of an agent are stored as part of its internal state, discernible at this LoA, adaptability results from the other two conditions. See L. Floridi, J. Sanders, *On the Morality of Artificial Agents*, in «Minds and Machines», 14, 2004, pp. 349-379.

⁶² Ivi., Allen et al. have proposed and analyzed the possibility of a *Moral Turing Test*. Allen, C. et. al., *Prolegomena to any future artificial moral agent*, «Journal of Experimental & Theoretical Artificial Intelligence», July 2000.

croethics of computers” or a simple deontology the user/programmer. An ethical theory that has profited from the combined anthropological and informational critique of the subject must be able to configure itself as patient-oriented, capable of surpassing in inclusivity – finding in information the minimal ontic and ontological level for moral consideration⁶³ – *non-standard* ethics (animal ethics, environmental ethics, etc.), overcoming the anthropocentric bias of classical macro-ethical theories, making the discussion of distributed morality around the actions of human, non-human, and “hybrid” agents in the world, online space, and infosphere more accessible. *Artificial Agents* (AAs), although obviously lacking *mens rea*, are capable of committing an *actus reus*⁶⁴ as, at the aforementioned LoA, they are capable of morally qualifiable actions. Producing actions that cause moral good or evil, thus moral actions⁶⁵. If one asks, like in the essay by Daniel Dennett, who

⁶³ Information Ethics raises information, as a common abstraction level for the consideration of every entity (which is at a minimal ontological level, non-contradictory and denotable), to a true patient subject of ethics. This allows overcoming the biases and anthropocentric impasses of standard ethics (think, for example, of the difficulties in universalizing Kantian imperatives) and the biocentric limit of non-standard ethics such as environmental ethics. Cfr. L. Floridi, *L'etica dell'informazione. Il fondamento filosofico dell'etica informatica*, in *Infosfera. Etica e filosofia nell'età dell'informazione*, Giappichelli, Torino 2009, Ch. I, pp. 25-57. Information Ethics configured in this way is based on four fundamental rules/principles: «(1) Entropy ought not to be caused in the infosphere (null law); (2) Entropy ought to be prevented in the infosphere; (3) Entropy ought to be removed from the infosphere; (4) The flourishing of informational entities as well as of the whole infosphere ought to be promoted by preserving, cultivating, and enriching their properties». L. Floridi, *Foundations of Information Ethics*, in ed. K.E Himma, H.T. Tavani, *The Handbook Of Information And Computer Ethics*, John Wiley & Sons, Inc., Hoboken, New Jersey 2009, Chapter I. pp. 3-20. This approach allows us to escape both the narrow confines of a “conservative” interpretation where the problems posed by the digital can be solved using traditional ethical tools, and those of isolation from the discussion and mainstream ethical methods due to a radical approach to information ethics as a completely new branch, bringing radically different problems from the past and necessitating completely ad hoc solutions. Cfr. L. Floridi, *Il male artificiale nella fondazione dell'etica informatica*, cit., pp. 73-76.

⁶⁴ Cfr. L. Floridi, *The Ethics of Artificial Intelligence*, cit., p. 185.

⁶⁵ Cfr. L. Floridi, J. Sanders, *On the Morality of Artificial Agents*, cit.

is to blame if the iconic evil AI of the Stanley Kubrick's movie *2001: A Space Odyssey* HAL 9000 kills someone⁶⁶.

«When HAL kills, who is guilty?», Floridi and Sanders state that «our approach allows us to say that HAL is accountable – although not responsible – if it meets the conditions defining action»⁶⁷. Imagine a patient asking for medical advice online and the response received saves their life. Does it matter if they received the decisive feedback for their health from a real doctor or an open-source chatbot? This extension of moral categories allows for renewing the definition of sociality, and thus politics, inherited and hypostatized from tradition. Wiener, the father of cybernetics, dealing with the moral consequences of the development of autonomous machines, stated that:

Society can only be understood through the study of messages and the communication media related to them; and that in the future development of these messages and communication media, messages between man and machines, between machines and man, and between machines and machines are destined to play an increasingly important part⁶⁸.

The key to understanding society is communication, which does not involve a monopoly of *Homo Sapiens*. The digital restructures ontologically and epistemologically the contemporary society. Now is necessary for non-human, artificial, and hybrid agents to be included and represented in their various capacities. As Wiener had foreseen, machines have joined with humans as active participants in society⁶⁹.

Thus, arises the urgency to rethink the very notion of political association, and here anthropological analysis can come in our help, because the age of the Fourth Revolution is the age in which we must

⁶⁶ D. C. Dennett, *When HAL Kills, Who's to Blame? Computer Ethics*, in *In Hal's Legacy: 2001's Computer as Dream and Reality*, edited by D. Stork, MIT Press, Cambridge MA 1997. pp. 351-365.

⁶⁷ «Quando HAL uccide, chi è colpevole? In realtà il nostro approccio ci permette di dire che HAL è imputabile – sebbene non responsabile – se soddisfano le condizioni che definiscono l'agire» (L. Floridi, *La moralità degli agent artificiali*, cit., p. 130).

⁶⁸ N. Wiener, *The human use of human beings*, Boston, Houghton Mifflin, p. 9.

⁶⁹ Cfr. T.W. Bynum, *Filosofia e Rivoluzione dell'Informazione*. In L. Floridi, *Info-sfera. Etica e filosofia nell'età dell'informazione*, cit., p. 11.

abandon the Constitution of the Moderns and its Parliament full of Modern Subjects. It's time to extend our democracies, and the public control that comes along with a political representation thought anew, to everything, machines, non-humans, hybrids, that has the features necessary to be a moral, social and political agent. «It is time, perhaps, to speak of democracy again, but of a democracy extended to things themselves»⁷⁰.

⁷⁰ B. Latour, *We have never been modern*, cit., p. 143.