The Future of AI
Extinction or Death of the Human?

PAOLO PELLECCHIA

Francisco de Vitoria University
paolo.pellecchia@ufv.es
ORCID: 0000-0002-5518-9133

Abstract: On March 30th, 2023, the CAIS released a statement warning against the danger of human extinction by the 'hands' of AI. By reclaiming the phenomenological centrality of death for human existence, I argue that this statement allows for an ontological diagnosis of both the human posture before the emergent progress of AI and the nature of this technology. To develop this diagnosis, I address three questions: What is the human? What is AI? What is the nature of the threat that AI poses to humanity? By drawing from Cassirer, Heidegger, Jankélévitch, and Landsberg, I firstly consider the human as *animal symbolicum moritūrus* and argue that the symbolic encounter with dead *loved* Thou represents a phenomenon of genuine human individuation. In following Dupuy, I then identify the core ontological feature of AI with model making and show how, in time, the model has taken ontological precedence over phenomenal. Lastly, I explain how the threat posed by AI coincides with the modeling abstraction of the *animal symbolicum moritūrus* into a dephysicalized *inforg*. Here, I maintain that the digitalized existence instantiated by AI coincides with an ontological shift: life reduced to an enhanced computational synthesis of matrixes is tantamount to an existence insolubly diverging from human life. AI progress is now forcing humanity to readdress the scandal that lays at the core of existence: not the extermination brought about by a Skynet-type AI, but the overt presence of a personalized death. What is the future going to look like? It is hard to tell, but I suggest that what awaits humanity is a grim fate were death not to take back its due ontological right.

Keywords: artificial intelligence, human existence, models, inforg, symbolic forms.

introduction

On March 30th, 2023, the Center for AI Safety (CAIS) issued a warning statement signed by a plethora of eminent scholars and AI scientist: "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war" (CAIS 2023). While the kind of extinction one should expect is left unclear, philosophical reflections that have so far addressed risks associated to AI mostly belong to the precincts of ethics (Lumbreras and Garrido-Merchán 2024; Hendrycks et al. 2023; Hutson et al. 2023; Hutson and Ratican 2023; Mazzi and Floridi 2023; Floridi 2023; DiMatteo, et al. 2022; Dubber et al. 2020, Green 2018). Without delegitimizing ethical methodologies, I follow instead the lesser treaded path of an onto-anthropological perspective and suggest that the warning should be taken seriously, not least because AI systems are inherently opaque, unintelligible, and uninterpretable to human understanding as they further complexify their internal structure and performance (Konig et al. 2022). Far from exaggerating, the short statement gives voice to genuine concerns and calls for a necessary diagnosis that must precede any political, ethical, or technical solution. To elaborate this diagnosis, I offer necessarily compressed answers to three fundamental questions implicitly solicited by the CAIS warning: If human existence is at risk, what is human existence? If AI represents a threat to the human, what is AI? What is the nature of the threat posed by AI?

1. What is Human Existence?

To tackle the first question, I draw from Cassirer's *An Essay on Man*. There is not room here to elucidate my choice vis-à-vis other interpretations of the *Anthropos*; suffices it to say that, by emphasizing the lived phenomenon of human existence, rather than inflexible metaphysical claims about human nature, Cassirer's perspective lends itself to a more fruitful dialogue with positions that, having emerged from traditional metaphysics, easily call

into question more rigid anthropological definitions. To Cassirer, the human is primarily a "strange mixture of being and nonbeing," a lived condition rife with incongruities that "traditional logic and metaphysics are in no [...] position to understand and solve [as] their first and supreme law is the law of contradiction" (2021, 11). As he argues, "the philosopher is not permitted to construct an artificial man; he must describe a real one. All so-called definitions of man are nothing but airy speculations so long as they are not based upon and confirmed by our experience of man" (2021, 11).

Notwithstanding the focus on experiential man, Cassirer does not resort to a definition in "empirical terms the way physical anthropologist and psychologist do" (Krois 2005, 558). Being concerned with a homogenous "*idée maîtresse*," empirical theories only provide us with facts whose very principle of selection depends on already present assumptions and, hence, pre-conceived understandings of the very object they try to define (Cassirer 2021, 21). Following Uexküll's theoretical biology, instead, Cassirer argues that "Man has [...] discovered a new method of adapting himself to his environment" (2021, 24): between *Merknetz* – a receptor system – and *Wirknetz* – an effector system – there is a region of experience that specifically defines lived human existence, namely a symbolic system, which mediates between the organism's biological receptivity and activity, individual and social activities. In Cassirer's words, "man lives not merely in a broader reality; he lives [...] in a new *dimension* of reality. [...] No longer in a merely physical universe, man lives in a symbolic universe" (2021, 24–25). Cassirer aims at correcting both empiricists and metaphysicians, not because reason does not belong to human experience, but on the account that a definition of man as a rational animal represents rather "a fundamental moral imperative:" "Reason is a very inadequate term with which to comprehend the forms of man's cultural life in all their richness and variety. But all these forms are symbolic forms. Hence, instead of defining man as an *animal rationale*, we should define him as an *animal symbolicum*. By so doing we can designate his specific difference" (Cassirer 2021, 26). In some unpublished manuscripts, Cassirer expands on this argument, expounding the doctrine of *Basisphänomene*, which claims the existence of three independent

and irreducible realities referred to by the pronouns I, Thou, It: "as he put it: 'Knowledge about 'me' is not prior and independent of knowing about 'You' and 'It', rather, all this is only constituted together" (Krois 2005, 562). In other words, "there is no *consciousness of a me* without consciousness of a you and even less is there a self, an 'ipse,' except in the general *Medium* of cultural forms, which provide the ways in which we are able to become a self" (Krois 2005, 564). Then, not only are the symbolic forms man is thrown into expressions of himself as *animal symbolicum* but they also represent principles of individuation: it is the relational character of the symbolic structure of human existence that individuates the human as a human.

I would like to radicalize Cassirer's formulation of man as *animal symbolicum* by adding the ontological and existential qualifier of *moritūrus*. That "*mors es hominis natura, no poena*" was already stipulated in antiquity (Pieper 2000, 64); even so, ontological mortality was generally coupled precisely with the rational component that Cassirer criticizes as insufficient for a suitable description of man. Here, instead, I bring the ontological weight of death together with an emphasis on the symbol-forming character of the human. However, if one must privilege the real man as postulated by Cassirer and make sense of the contradictory nature of death insofar as it imbues an incongruous human existence, then a metaphysical or empirical anthropology will be to no avail (Rahner 1961, 54). In this sense, existential phenomenology seems to be the most adequate methodology to capture death's foundational import for the ontology of human existence. I mainly draw from Heidegger, Jankélévitch, and Landsberg for their perspective allows to also integrate the fundamentally relational character of the *animal symbolicum* revealed by the doctrine of *Basisphänomene*.

As it is known, Heidegger holds that "death is the possibility of the absolute impossibility of Dasein. [...] This ownmost possibility, however, non-relational and not to be outstripped, is not one which Dasein procures for itself subsequently and occasionally in the course of its Being. [I]f Dasein exists, it has already been *thrown* into this possibility" (1962, 294). For Heidegger then mortality coincides with "an essential distinguishing mark of what it means to be human" (Wrathall 2005,

62). Although "only in dying can I to some extent say absolutely, 'I am,'" (Heidegger 1985, 316–318), Dasein cannot experience the event of its own death nor the one of the other (Heidegger 1962, 281ff.): at best, Dasein can be there alongside the dying of the other. Through the non-relational character of death, Dasein is individuated for its ownmost potentiality discloses to it when it resolutely confronts the constantly imminent possibility of death face-to-face. However, in experiencing an anxiously "impassioned freedom towards death" (Heidegger 1962, 311), Dasein is also individualized for it discloses to itself as unrelated, "*unsupported by concernful solicitude*" (Heidegger 1962, 311), stripped away of the everydayness of the They-world. But is that so? In *La mort*, Jankélévitch maintains that death is never entirely an a-tragic problem (an external event that only a third person suffers) nor an a-problematic tragedy (an event that only a first person undergoes in its isolated being-towards-death) (1977, 31). While normally oscillating between these two apprehensions of death, man is occasionally enabled to dwell in the frictional tension of a "*quasi mortem propriam*" (Jankélévitch 1977, 30) by the death of the Thou. This represents the deictic event where "the universal law of mortality is experienced as [...] a personal tragedy; and [...] a necessity of general order" (Jankélévitch 1977, 31). In *La mort*, however, the necessary isolation of the dying person and the ecumenical character of death never resolve themselves into one another but, rather, constantly push back against each other.

By arguing that the uniqueness of the other reaches a degree of heightened tension with the death of a *loved* Thou, Landsberg goes beyond both Heidegger's non-relational character of death as principle of individuation and Jankélévitch's quasi-identification with the dying Thou. It is here that death is no longer either a simply present metaphysical possibility of the utmost impossibility or a paradoxical experience defined by a minimal yet unescapable distance. With the death of a friend, "we meet death in its own personal uniqueness. Each death is unique as each person's manner of being present" (Landsberg 1953, 13). Authentic individuation, to Landsberg, arises only from the intuition of the necessity of my own death that springs forth from the death of a *loved* Thou with whom I establish a participative community; within

the 'we' we constitute with the dying man "for a moment we have our feet in the land of the dead" (Landsberg 1953, 14). The direct experience of one's personal death as a principle of individuation and selfhood is then symbolically mediated by the death of the *loved* Thou. Through this symbolic and affective mediation, I am ecstatically called out from the biological flux of my natural demise and dragged face-to-face if not with my actual death, at least with a foretaste of the highest density: here, I am gripped by the intuitional realization that my ontologically authenticity coincides with the fact that not only am I "moribundus" (Heidegger 1985, 317), but *moritūrus*. Simultaneously, I am also gripped by the experiential event of an unthinkable communion with the dead; this mysteriously symbolic society reveals an onto-anthropological dependence of the I on the dead Thou. As Cassirer notes, man is an *animal symbolicum* that attains consciousness and selfhood through the mediation of cultural forms by means of which the society of the human expands to the dead: they become part of the individuating *locus* whence man can say 'I.' As I will argued later, these symbolic forms are authored by the dead and show the *animal symbolicum* that its authentic individuation coincides with its condition of *moritūrus*. This condition ec-statically historicizes man, taking him out of the natural flux and into the symbolic worldhood of culture. As Harrison writes, "what intervenes in natural time is human finitude, which is unlike other finite things in that death claims our awareness before it claims our lives. We dwell in space, [...] but we dwell first and foremost within the limits of our mortality" (2003, 19). It is by means of mortality, by the individuation spawned through the dead-*loved* Thou, that man can build a meaningful existence through the institution of symbolic forms. These are the place where the dead "indwell" (Harrison 2003, x), the site of their dominion and helplessness, the site of human inheritance and responsibility.

2. What is AI?

The question about the essence of AI is a thorny one and a consensus in this matter is fundamentally lacking nowadays. Sheikh et al. argue that

the contrast between the abundance and volatility of definitions and the paucity of an ontological grasp is not due "to carelessness but is inherent in the phenomenon of AI" (2023, 15) whose very heterogeneity makes a shared definition unattainable (König 2022, 23). Indeed, Wang notes that "the current field of AI is actually a mixture of multiple research fields, each with its own goal, methods, applicable situations, etc., and they are all called 'AI' mainly for historical, rather than theoretical, reasons" (2019, 28). The historical character of this nomenclature, moreover, shows that the discipline is subject to constant change over time (Bartneck et al. 2021, 8), thus "making AI a moving target" (König 2022, 24) and its definitions useless (Wang 2019, 5). Given the slippery definitional character of AI, I propose to follow Dupuy's archeological approach to expose AI's core ontological traits and hopefully arrive at an epistemological clarity that seems hard to attain when approaching AI merely from a technical viewpoint.

Dupuy holds that cognitive science reclaimed for itself questions that traditionally belonged to the precincts of philosophy and theology, specifically the mind-body problem and the hard problem of consciousness (2009, 90). Cybernetics was no less ambitious when McCulloch maintained that organisms are not merely "analogous to machines but that they are machines. [and] brains are a very ill-understood variety of computing machines" (1965, 163). Surely, McCulloch's position does not represent the only viewpoint held by cyberneticians; yet, it reveals that cybernetics too was characterized by radical ontological claims right from its inception. Such claims were foundational of both the outlook that pushed forward cybernetic endeavors and the philosophical inheritance left to those fields that took over cybernetic's ontological assumptions, artificial intelligence being one of these (Dupuy 2009, 57). In a way, both cognitive science and cybernetics are defined by the paradox of an "enterprise that sets itself the task of *naturalizing* the mind [and to do so holds] as its spearhead a discipline that calls itself *artificial* intelligence" (Dupuy 2009, 21).

To advance their ontological avowal, cognitive science and cybernetics used primarily one tool: the model. For Dupuy "a model is an idealization, usually formalized in mathematical terms, that synthesizes a system of relations among 'elements whose identity and even nature are, up to

a certain point, a matter of indifference, and which can [...] be changed, replaced by other elements, analogous or not, without [the model's] being altered'" (2009, 29). By abstracting a highly detailed phenomenal reality, a model simplifies it by selecting recursive patterns and assembling them into a system of functional relations that reproduce an explanatory and predictable simulation of the real object. As Wiener writes, "no substantial part of the universe is so simple that it can be grasped and controlled without abstraction. Abstraction consists in replacing the part of the universe under consideration by a model of similar but simpler structure. Models [...] are thus a central necessity of scientific procedure" (1945, 316). Such simplification of reality is more "readily mastered" and thus runs the risk to become the sole center of scientists' attention (Dupuy 2009, 31). Since "the intention and the result of a scientific inquiry is to obtain an understanding and a control of some part of the universe" (Wiener 1945, 316), achieving such a feat becomes tantamount to creating models that follow the procedural operations of "reproduction, representation, repetition, simulation" (Dupuy 2009, 32). These steps are part and parcel of the scientific method since Bacon, but cognitive science turned them into the only epistemological path for knowledge acquisition (Dupuy 2009, 32).

Among these steps, simulation becomes the privileged theoretical tool to realize a model that ideally relates to reality according to a one-to-one correspondence. Following Turing's work, the underlying isomorphic creed that defines model making (the belief that abstractable traits coincide with essential patterns that repeat themselves in different domains of experience) develops into a logical equivalency between models and reality. This creed led to the conclusion that "the mind is a physically constructible machine and therefore replicable, perhaps, by human artifice" (Dupuy 2009, 53). In 1943, McCulloch and Pitts pushed the epistemological value of isomorphism to the biological level: now, not only the mind, but also the biological structure of the brain can be effectively simulated and grasped by an equivalent machine. It is not hard to see why AI is downstream from cybernetics. It is worth adding that the isomorphic equivalency predicated on simulation, at some point, shifted toward actual replacement, where those patterns that have been pursued, abstracted, and encapsulated in the model become the very

essence of phenomenal reality. Due to the work of McCulloch and Wiener (just to provide two honorable mentions), the very difference between phenomenal reality and its description (achieved via reproduction, representation, repetition, simulation) begins to vanish. Reality then becomes "a mere means to a higher end: model making" (Dupuy 2009, 138). Under the sway of cybernetics, the artificial ontologically supplants the natural while the natural, in turn, is subsumed into the ontological status of the artificial, becoming fully knowable and explainable insofar as its processes can be modelled and simulated through human artifice. One may argue, then, that both the ontological overlapping just noted and the precedence of the artificial over the natural lay the ground for what Oderberg defines as telic possibility. In the case of cybernetics, it does not really matter if one experiences phenomenal reality or its simulation; what matters is that the artificial shows patterns, produces effects, and displays purposes (Oderberg 2012, 385) that are good enough to replace the natural.

The ontological attitude of cognitive science and cybernetics represents the philosophical bequest passed on to artificial intelligence: indeed, "the very concern for universality exhibited by artificial intelligence makes it possible to regard this enterprise […] as a philosophy – and a philosophy of a transcendental type at that" (2009, 93). Given such inheritance, one may argue that what Breton observes about cybernetics be also applied to AI: "Cybernetics […] assumes [the form of] a terrible paradox: it affirms humanity while at the same time depriving man of it. In this sense, it rather openly expresses a fundamental characteristic of contemporary technical and scientific knowledge, by virtue of which the benefits of progress seem irrevocably associated with the rational portrayal of the death of man" (1984, 160). This death, however, is an onto-anthropological one: who becomes extinct is not the human as a biological species but man as *animal symbolicum moritūrus*. Then, one can perhaps give another try to a definition of AI: if Heidegger (1976) defined cybernetics as the metaphysics of the atomic age and Dupuy characterizes it as "a decisive step in the rise of *antihumanism*" (2009, x), AI, as cybernetics' most faithful legatee, could be seen as the metaphysics of the post-human age.

3. What is the Threat of AI?

In 2017, Jang Ji-sung's daughter Nayeon died of haemochromatosis. They were reunited two years later in a virtual park through a VR headset. By the end of the encounter, "Nayeon runs and plays, and tells Ji-sung she isn't sick anymore. [Then] Nayeon lies down to sleep and morphs into a butterfly" (Stokes 2021, 130). A virtual burial of some sort took place: a somehow irresponsive, yet sufficiently interactable AI simulation of a dead daughter was able to bring comfort to a distraught mother. The telic logic underlying this event is evident: Nayeon's AI replica was good enough to remedy the irreconcilable tragedy of her own death. To Stokes, the most concerning aspect of AI replicas is of ethical nature: on the one hand, the replacement reduces the dead "to a mere resource for fulfilling our needs" (2021, 141); on the other, the assumption that an AI avatar could take over when a person is no longer alive deprives the living of their sacrosanct irreplaceability, turning them into "fungible goods" (Stokes 2021, 143). Both the dead and the living become replaceable. While Stokes is right, I argue that the root problem should be seen in the onto-anthropological shift that these technologies impose on the *animal symbolicum moritūrus*.

To explain the nature of the threat posed to the human by AI, one should understand that the autonomy of the model and the ontological precedence it acquires over phenomenal reality are rooted in the emphasis that cybernetics put on the "question of *information*" (Dupuy 2009, 113): if an intellectual construction can, or perhaps should, replace the real object it is because they embed the same amount of information due to an ontology-laden isomorphism. Floridi argues that the primary aim of Information and Communication Technologies (ICTs) is a philosophical one: by changing "our understanding of the world and of our interaction with it, [as well as] in our self-assessment and identity" (2010, 8), ICTs re-ontologize our world (2010, 11). He holds, then, that "one might interpret reality as constituted by information, that is, by mind-independent, structural entities that are cohering clusters of data, understood as concrete, relational points of lack of uniformity" (2010,

71), and maintains that the ultimate nature of reality consequently morphs from a materialistic to an informational one. This shift coincides with a fundamental transformation of objects and processes into de-physicalized, typified, and interactional entities: object and processes lose their unrepeatability and turn into "perfectly cloneable" types, whose criterion of existence lies in their "being potentially subject to interaction" (Floridi 2010, 12). To be, now, means "to be interactable" (Floridi 2010, 12). Due to this informational shift, man is re-ontologized into the *inforg*: an interconnected informational organism (Floridi 2010, 9) that inhabits the new republic manufactured by the informational revolution as a de-physicalized, typified, and interactable citizen, exhibiting "the intrinsically informational nature of human agents" (Floridi 2010, 10). Is Nayeon any different?

To rectify the abstracting informational re-ontologization of the *animal symbolicum moritūrus* into the post-human *inforg*, I propose to align with Harrison's recuperation of the "humic foundation of our life worlds" (2003, x) as they take shape into symbolic forms. These forms historicize and give a concrete existence to man as *animal symbolicum moritūrus* and his activity. But who authors the symbolic forms created by human activity and takes possession of human life and conduct? I contend that such aboriginal founders are the dead who beget human symbolic existence and represent the foundational principle of individuation. Man comes into being by being placed into a pre-existing world of symbolic practices and stays into being by actively partaking in and shaping the very symbolic practices that precede him: man's being is always mediated by his placeness and by the placing of himself into the chronotopic site whence his existence is concretely grounded. As Harrison notes, "if human being is a 'being in the world,' […] then nothing is more essentially related to our being than our building practices, given that the world into which we are thrown is always a *built* world. We take our measure of being from what surrounds us; and what surrounds us is always, to some extent, of our own making. Building, which starts from the ground up, is where the fundamental ontology of our mundane lives both begins and ends" (2003, 17). Then, again, who speaks through a place is our forefathers, our dead founders.

But what is a place? To Harrison, a place is the site where "human worldhood" (2003, 19) meaningfully domesticates space, affording it with symbolic value. Human worldhood, however, is also temporally defined: a "place is where time, in its human modes, takes place" (Harrison 2003, 19). As argued earlier, the specific human mode of existing in time is that of finitude and mortality, so that, "when we build something in nature [...] we create the rudiments of a world and thereby give a sign of our mortal sojourn on the earth" (Harrison 2003, 19). To Harrison, the most archetypical expression of human mortality and finitude is the institution of burial in that it founds civil society by binding "the living to the dead and the dead to the unborn" (2003, 81). Through the symbolic mediation of burial death, as principle of individuation, takes on the concrete face of the dead, that is, the familiar face of the *loved* Thou. The site of the burial is the *locus* wherein man is given worldhood and subjectivity; but it is also the domesticated site whence human worldhood and subjectivity shape space into a place, affording nature with a symbolic meaning, because they are firstly shaped by the authority of the dead that lay therein.

In a way, then, every symbolic form is an act of burial. Not simply in the sense of laying the dead to rest in the ground, but with the meaning of storing and preserving (Harrison 2003, xi), of thickening phenomenal reality with world-making actions that stow a latent signification in it. This burial act always leaves behind an inscription: *hic jacet*. Harrison maintains that "the *hic* of *hic jacet* is the aboriginal *Da* that grounds *Dasein*'s situatedness and historicizes its being in the world, especially since *jacet* alludes to the finite temporality that *Dasein* makes its own in its so-called being-toward-death" (2003, 22). If this is true, then, what is lost with the transformation of the human into an *inforg* is precisely this situatedness: de-physicalized and typified, the *inforg* is uprooted and placeless, nor here or there, but everywhere and hence nowhere. The *inforg* is groundless in so far as it is a simplified abstraction, a typification of the human which has been elevated from the earthly ground of existence into the informational spheres of the infosphere. In this sense, one may say that the virtual burial of Nayeon did not take place. The green screen that harbored the farewell between a grief-stricken mother and her simulated daughter was neither green nor greener pastures; it was, indeed, of

a conspicuously different order than space domesticated into a place by human mortality. There, no corpse indicated the removal of the self, nor any *Da* pointed "to that persistent finitude that underlies the placehood" that pervades Dasein (Harrison 2003, 22). On that green screen, no *hic jacet* was chiseled; under that space, no corpse was lying latent, deeper than the shallow bidimensionality of a cluster of data synthetized by a binary code. In this sense, Nayeon's virtual burial was deprived of the symbolic thickness and transcendence that guarantee any relation with the dead. Without casting any judgment, the fact that Nayeon's replica morphs into a butterfly instead of sinking into the ground is indicative of the uprootedness that defines the flattened garden where the relationship between Nayeon and her mother dissolved rather than being rescued.

The removal of the symbolic dimension of the burial implies the loss of both the foundational relationship with the dead and the possibility to be individualized through the death of the *loved* Thou. The transformation of the *animal symbolicum moritūrus* into an *inforg*, then, results in the eradication of both the attributes *symbolicum* and *moritūrus*. The virtualization of the burial thus alludes to an anthropological reduction in two senses. On the one hand, it coincides with a simplification according to Wiener's definition of a model: the contradictory heterogeneity of human existence is simplified and typified into a "*mathesis universalis*" (Cassirer 2021, 49), that is, a logically defined operational praxis that employs a computational model as a translating process. The phenomenal *animal symbolicum moritūrus* is literally translated into the bidimensional binary language of coding. On the other, this simplified translation eradicates the constitutive relational dimension with the Thou that the symbol engenders. The loss of the symbolic value of burial, understood as the site where the foundational relationship with the concrete death of a *loved* Thou takes place, implies a metamorphosis of those properly human relationships that are fostered and safeguarded by the symbol. Due to this loss, human relations are uprooted and subsumed into technical interactions (Sheehan 1981, 56). As a result, it is no longer Cassirer's real man who lives and dies on the earth through his symbol-making active existence, but rather the impalpable *inforg* that soars through the infosphere. By being de-physicalized, the *animal symbolicum moritūrus*

is taken away from his concrete situatedness, stripped away from the place wherein he exists as a human, literally displaced and deprived of the meaning-laden place where he communes with the dead. Here, the subject as *subjectum* disappears because there is no longer anyone who is literally 'thrown under,' anyone who, by inhabiting the earth from within, can provide the symbolic foundations on which human existence and culture are built.

Nayeon's virtual burial is but an exemplifying hypostatization of the threat posed by AI to humanity in that it occasioned the transformation of her unique Thou into a de-physicalized and typified cluster of data: the real Nayeon, who nonsensically died, was immortalized into an *inforg*. If it is true that death wears "the mask of the dead" (Harrison 2003, 93), Nayeon's morphing into a butterfly signifies the shattering of the mask, the covering-up of death as principle of individuation. If it is true, as seen so far, that death must "proceed from the corpse" (Harrison 2003, 92) of the *loved* Thou, the dematerialization, informationalization, and 'dataization' of the human implies an impoverishment of the experience of death for the subject and a consequent impoverishment of selfhood for both the living and the dead. The loss of the body for the living and of the corpse for the dead fabricates the illusion that death, no longer ontologically foundational of the human mode of inhabiting existence, can be overcome. But a world without humans that die, without mortals, is a world of beings that have been uprooted, unrooted, dehistoricized, and, hence, dehumanized. It is in this sense that digital immortals are un-human: being de-humanized, deprived of their humic roots, of the authority of the dead and of death, they retain nothing of human existence. In the storing of data that builds humanity's new make-up there is nothing that is really stored: nothing of the unique person that has been digitalized is preserved because nothing of the unique death of the person can be preserved. The replacement of the *loved* Thou for an AI replica is tantamount to a silencing of the dead and their dominion, but this silence is nothing different than a form of displacement that affects the living: no longer engaged in a *convivium* with the dead through the symbolic forms, they too are displaced from their human condition.

Conclusions

The paradigm change effected by cybernetics and pursued by AI may very well represent an onto-anthropological threat before an ethical one: by turning the *animal symbolicum moritūrus* into an *inforg*, AI affords humanity with a subjectless ontology, "announcing the dehumanization of man" (Dupuy 2009, 158). If "humanity is not a species [but] a way of being mortal and relating to the dead" (Harrison 2003, xi, 34), then the de-symbolization and artificial immortalization of humanity, along with the disappearance of the *subjectium* produced by AI, might mean the end of the human condition. Informatized and datasized, that is unearthed and subjectless, the human is no longer able to die, nor to symbolically relate to phenomenal reality and its root-foundations. Then, perhaps, the statement issued by CAIS should not be looked down upon as a sign of alarmism but as a description of an approaching disappearance. What is in danger is not the human species; we will continue to die, perhaps. What is in danger is a specific human mode of being in the world, that of the *animal symbolicum moritūrus*.

Split between attraction and fear, human beings are now compelled to readdress the central scandal that lays at the core of existence and that AI has forced to face again: not the death we are vaguely conscious about as we go by life, nor an extinction brought about by a Skynet-type AI. Rather, man is called to face again the overt presence of a personalized death that abides within one's own life. In this situation, "each of us must choose an allegiance – either to the posthuman, the virtual and the synthetic, or to the earth, the real and the dead in their humic densities" (Harrison 2003, 34). What is the future going to look like? It is hard to tell, given the nature, evolutionary pace, and decentralized domain of current AI models. I would suggest, however, that what awaits humanity is a grim fate were death and the dead not to take back their due ontological right. Here, the telic ontology of Nayeon's burial will not be good enough to rescue humanity from its fate.

Bibliography

Bartneck, Christoph et al. 2021. *An Introduction to Ethics in Robotics and AI*. Cham: Springer.

Breton, Philippe. 1984. "La cybernetique et les ingenieurs dans les annees cinquante." *Culture technique* 12: 155–161.

CAIS, 2023. "Statement on AI Risk. AI Experts and Public Figures Express their Concern about AI Risk." Accessed July 31. https://www.safe.ai/work/statement-on-ai-risk.

Cassirer, Ernst. 2021. *An Essay on Man*. New Haven: Yale University Press.

DiMatteo, Larry A., Cristina Poncibò, and Michel Cannarsa (eds.) 2022. *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press.

Dubber, Markus D., Frank Pasquale, and Sunit Das. 2020. *The Oxford Handbook of Ethics of AI*. Oxford: Oxford University Press.

Dupuy, Jean-Pierre. 2009. *On the Origins of Cognitive Science. The Mechanization of the Mind*. Cambridge: The MIT Press.

Floridi, Luciano. 2023. *The Ethics of Artificial Intelligence. Principles, Challenges, and Opportunities*. Oxford: Oxford University Press.

Floridi, Luciano. 2010. *Information. A Very Short Introduction*. Oxford: Oxford University Press.

Green, Brian Patrick. 2018. "Ethical Reflections on Artificial Intelligence." *Scientia et Fides* 6(2): 9–31. DOI: http://dx.doi.org/10.12775/SetF.2018.015.

Harrison, Robert P. 2003. *The Dominion of the Dead*. Chicago: The University of Chicago Press.

Heidegger, Martin. 1985. *History of the Concept of Time: Prolegomena*. Translated by Theodore Kiesel. Bloomington: Indiana University Press.

Heidegger, Martin. 1976. "Nur noch ein Gott kan uns retten." *Der Spiegel* 23: 193–219.

Heidegger, Martin. 1962. *Being and Time*. Translated by John Macquarrie and Edward S. Robinson. Oxford: Blackwell.

Hendrycks, Dan, Mantas Mazeika, and Thomas Woodside. "An Overview of Catastrophic AI Risk." Accessed July 31, 2024. https://arxiv.org/pdf/2306.12001.

Hutson, James, Jeremiah Ratican, and Colleen Biri. 2023. "Essence as Algorithm: Public Perceptions of AI-Powered Avatars of Real People." *DS Journal of Artificial Intelligence and Robotics* 1(2): 1–14

Hutson, James, and Jeremiah Ratican. 2023. "Life, death, and AI: Exploring digital necromancy in popular culture. Ethical considerations, technological limitations, and the pet cemetery conundrum." *Metaverse* 4(1): 1–12.

Jankélévitch, Vladimir. *La mort*. Paris: Flammarion, 1977.

König, Pascal D. 2022. "Essence of AI. What is AI?" In *The Cambridge Handbook of Artificial Intelligence*, edited by Larry A. DiMatteo et al. Cambridge: Cambridge University Press.

Krois, Micheal J. 2005. "'A Passion Can Only be Overcome by a Stronger Passion:' Philosophical Anthropology before and after Ernst Cassirer." *European Review* 13(4): 557–575.

Landsberg, Paul-Luis. 1953. *The Experience of Death*. London: The Camelot Press.

Lumbreras, Sara and Eduardo Garrido-Merchán. 2024. "Insights from Saint Teresa and Saint Augustine on Artificial Intelligence: Discussing Human Interiority." *Scientia et Fides* 12(2): 265–295.

Mazzi, Francesca, and Luciano Floridi. 2023. *The Ethics of Artificial Intelligence for the Sustainable Development Goals*. Cham: Springer.

McCulloch, Warren S. 1965. *Embodiments of Mind*. Cambridge: The MIT Press.

Oderberg, David. 2012. "Disembodied Communication and Religious Experience: The Online Model." *Philosophy and Technology* 25: 381–397.

Pieper, Josef. 2000. *Death and Immortality*. South Bend, IN: St. Augustine's Press.

Rahner, Karl. 1961. *On the Theology of Death*. New York: Herder and Herder.

Sheehan, Thomas, ed. 1981. *Heidegger: The Man and the Thinker*. Chicago: Precedent.

Sheikh, Haroon, Corien Prins, and Erik Schrijvers. 2021. *Mission AI*. Cham: Springer.

Stokes, Patrick. 2021. *Digital Souls*. London: Bloomsbury.

Wang, Pei. 2019. "On Defining Artificial Intelligence." *Journal of Artificial General Intelligence* 10(2):1–37.

Wiener, Norbert. 1945. "The Role of Models in Science." *Philosophy of Science* 12(4): 316–321.

Wrathall, Mark. 2005. *How to Read Heidegger*. London: Granta.