New Techno-Humanities 1 (2021) 100001

Contents lists available at ScienceDirect

New Techno-Humanities

journal homepage: www.elsevier.com/locate/techum

# The Faustian Machine and the Chrome Lotus: On the diversity of perspectives on the metaphysics of artificial intelligence with a particular focus on the contributions of traditional non-Western thought

Dr. Johan Eddebo

*Centre for Multidisciplinary Research on Religion and Society, Uppsala University, Teologiska institutionen, Box 511, 751 20 Uppsala, Sweden*

ARTICLE INFO

*Keywords:*
Artificial intelligence
Metaphysics
Non-Western philosophy
Religion and science
Philosophical anthropology

ABSTRACT

This article intends to provide an outline of the diversity of philosophical perspecives pertaining to artificial subjectivity, with special regard to non-Western approaches and their potential to contribute to the discussion. The premise is that the popular-scientific as well as academic debate on artificial subjectivity is somewhat narrow, with a bias towards reductionist and functionalist approaches. This prevents the inclusion of the plethora of perspectives inherent in the Western philosophical tradition, while also blocking potentially fruitful input from non-Western traditions. The outsider role of the latter perspectives in relation to the narratives of modern technological society has the clear potential to supplant them where they become myopic, and this overview provides several examples of possibly fruitful contributions. Of particular importance is their evident sensitivity to the political and cultural implications of technology, as well as a set of unique methodologies and approaches suited to uncovering and addressing problems pertaining to the interactions between technology, individuals and society.

## Introduction

It is a fairly non-controversial statement that bodies are somehow associated with consciousness. In any case, we seem to have an immediate experience of the connection between our own phenomenal subjectivity and a physical body, however "physical body" is cashed out ontologically.

But we know quite little about how this connection comes to be, what consciousness as such really is, or how it originates. These metaphysical issues are indeed met with many plausible answers in various philosophical and religious traditions of knowledge, yet unless we definitively side with one of these, the situation is quite unclear.

In spite of this fact, however, the contemporary popular-scientific discourses pertaining to artificial intelligence often enthusiastically embrace the possibility of genuine artificial subjectivity as a matter of fact (see e.g. Pavlus 2019). There is scarcely a nod towards the often significant metaphysical difficulties inherent in such a proposition, and when there even is a discussion of the causal minutiae, it is generally assumed that the mind is a kind of software in our "brain computers" that we must be able to recreate upon any sufficiently suitable substrate.

The situation isn't much better within philosophy of mind as a discipline, which generally affirms the metaphysical possibility of artificial subjectivity as a matter of course. Some variant of reductionist functionalism akin to the popular view holds the consensus position here as well – that is, any analysis of the mental that attempts to reductively define it in terms of the activity of something akin to a computer program or as a rule-governed function. This position is currently supported by prominent names such as Metzinger and to some extent even Chalmers, in spite of Kripke's forceful rebuttal of its principles almost forty years ago which is yet to be defused (Buechner 2011, 342-369), not to speak of the incommensurability problems in the Nagel-Sprigge tradition that will undermine any reduction of the subjective to objective states of affairs (see Baker 2013 for a detailed discussion of this category of arguments).

Another significant problem is that classical philosophy and non-Western traditions of knowledge are very rarely allowed to bear upon this debate, which most often is framed by a scientistic discourse characterized by the presumptions of modern analytic philosophy, and almost exclusively draws on European traditions.

With the aim of not only fostering a cross-pollination of ideas, but also to address the arguably myopic character of anglophone philosophy, the intention of this article is therefore to nuance the current debate (Schwitzgebel et al. 2018, 21-48). I wish to present and provide an outline of the diversity of philosophical perspecives on this all too prominent set of issues that hopefully will serve to challenge one or two entrenched prejudices, with a particular focus on the non-Western approaches and their potential to contribute to the discussion.

In detail, I intend to provide an overview of the available mainline Western ontologies pertaining to the human mind, and assess their respective portrayals or conceptual framing of the metaphysical possibility

*E-mail address:* Johan.Eddebo@crs.uu.se

https://doi.org/10.1016/j.techum.2021.100001
Received 15 May 2021; Received in revised form 8 September 2021; Accepted 10 September 2021

of artificial consciousness. The focus will be on sketching the problems inherent to these positions in terms of affirming artificial consciousness. Secondarily, I will specify what unique prospects the different positions bring to the table in terms of addressing legacy problems of consciousness, and what I consider to be some of their general benefits. The following major part of the article focuses on non-Western philosophical traditions using similar criteria.

*Definitions of consciousness*

First of all, what is then this fact of consciousness around which the debate moves? What is the "explanandum" that we're trying to describe, if not exhaustively explain here? Ambiguities at this level will influence the entire succeding discussion and every argument in it, so it is of vital importance that the basic definitions are in some kind of order and do not implicitly beg the question in any direction.

Nonetheless, it is obvious that we need to start with the fact of phenomenal consciousness as such, however we are to frame it more specifically. That is, we have to start with subjective experience, with the "what-it-is-like" to *be* from the subjective point of view as an immediate fact, and then move on to see whether e.g. consciousness ultimately reduces to some other set of facts, or itself is metaphysically robust. In other words, we can't "define away" the problem as such because it pertains to an immediate fact of experience – but neither must we beg the question when we try to frame this fact discursively. If we for instance tend either towards a behaviourist or idealist description of the basic fact at hand, this will corrupt the entire discussion afterwards. In other words, the dualist cannot start off with describing consciousness as something like "the movement of the spirit, unperturbed by the vagaries of matter", and neither can the materialist begin by conceiving of subjectivity as "just the objective experience of a material being in a material world".

In other words, notwithstanding how consciousness can be drawn up or differentiated in detail, such as according to Lycan's typology, there is a bare, immediate phenomenal fact that must be addressed (Lycan 1996, 1-4).

*Definition of artificial consciousness*

With the above remarks in mind, an adequate definition of artificial consciousness can be formed.

Since we are within the framework of a discourse of technology, and must inevitably assume something like the opposition inherent in the artificial-natural dichotomy, we are going to have to exclude such things as natural procreation. Accordingly, we cannot define as artificial a form of consciousness which merely is the effect of a human cause in the most general sense. We need *artifice* in the picture, so the consciousness in question must be a product of our ingenuity and artistry, rather than our own (or any other's) natural progeny.

A basic definition satisfying these constraints would describe artificial consciousness as any form of first-person awareness that is not an attribute of the natural offspring of a conscious species.

Still, this leaves us with quite a few options in practice. Aside from the hypothetical technological substrate of chrome, microchips and software, alternatives such as genetically engineered biological matter would also qualify. The crossover variant of cells imbued with nanotechnology to manifest subjectivity is also an interesting alternative.

## The mainline metaphysical perspectives and their approaches

Having established the explanandum, we can now move on to examine the diversity of options for metaphysically framing the basic fact of consciousness and how an artificial one from the point of view of these options could be construed as metaphysically possible. This overview will be limited in its representation of perspectives outside of the historical mainstream of Western thought due to my own biases and restricted knowledge, but care will be taken to as far as possible affirm their unique character and not falsely equate them to similar points of view in the conventional philosophical discourse.

*Metaphysical idealism*

Metaphysical idealism can either be summarized as the hypothesis that everything can be reduced to some form of conscious experience, as per Berkeley (Berkeley 1996, 25-30), or that the ontological basis of reality is ultimately comprised by non-physical, abstract objects, in line with the idealistic conceptualizations of Platonic realism and their subsequent offshoots (Silverman 2014). The two options will relate to artificial consciousness in somewhat different ways.

The first alternative amounts to the position that there is no material substrate for consciousness in the literal sense. There is only consciousness. Affirming artificial consciousness in accordance with the above definition would then metaphysically speaking entail either the "channeling" of a pre-existing consciousness onto a pseudo-material substrate, or the production of a higher-level consciousness by combining lower-level ones – perhaps as a supervenient phenomenon.

The second alternative offers a broad range of options, depending on how it is circumscribed. It is compatible with the reductive idealist position just mentioned and the associated solutions, e.g. if each experiencing subject is considered a particular immaterial form that can neither be generated nor corrupted, but possibly conveyed. At the same time, something like Platonic realism could underwrite the synthetic generation of any subject, provided it is conceived as the embodiment or manifestation of an abstract universal.

*Cartesian dualism*

Cartesian dualism, the position that mind and body are discrete substances, offers us the same possibility of channeling a pre-existing subject in relation to a material substrate. It is also in principle compatible with a kind of sui generis production of consciousness, such as if the "correct" combination of matter simply causes an independent mental substance to emerge by virtue of some unknown psychophysical laws and causal principles.

*Panpsychism*

According to panpsychism, the computer already is conscious in some rudimentary sense. The position is akin to metaphysical idealism, and the two are to some extent compatible. Panpsychism would however add, that even if there really is something irreducibly material, that cup of coffee in your hand is nonetheless conscious, as in possessing some kind of primitive subjectivity. As with metaphysical idealism of the Berkeleyan kind, the position seems to conceptually allow for the synthetic generation of "higher-level" consciousnesses to be generated by combining less complex ones.

*Property dualism/emergentism/non-reductive physicalism*

In accordance with this set of positions, mental properties could in principle simpy supervene on the right type of physical substrate. Indeed, on property dualism and the associated accounts, all consciousness is really "artificial" on the above definition insofar as it really emerged from matter as such rather than due to the form or potentiality of any natural progenitor. The classical interaction problem therefore becomes a central obstacle to these perspectives, since the procession of an effect qualitatively distinct from its cause, i.e. one that is in no way found in the cause, is problematic in relation to the basic principle of causality.

Also, these positions are notoriously unstable and are often argued to either collapse into reductive physicalism or some immaterialist position due to delineation difficulties (e.g. panpsychism, metaphysical idealism or some equivalent to substance dualism), in which case we would either

have to include their options, or default back to one of them (Eddebo 2017, 70-73; 105-110).

*Hylomorphic dualism*

On this position, anchored in Aristotelian thought, the embodied form is what possesses subjectivity. Hylomorphism is probably the most complex of positions recounted here, since it balances a dualist outlook with a view on single substances that irreducibly fuses together the material and the non-material. It has parallels to several of the preceding positions.

Firstly, it strictly distinguishes between animal and human subjectivity, akin to Lynne Baker's distinction between primitive and robust types consciousness (Baker 2013). The human consciousness has both material and immaterial parts, while the more primitive kinds are exclusively material. However, the hylomorphist's conception of matter has an affinity to panpsychism. On this position, all bodies are necessarily the metaphysically (but not logically) irreducible union between the immaterial and the physical aspects of reality, so there is something comparable to subjectivity in every "informed" physical body, which resides in the form. Forms cannot be generated, so we have the same basic situation as with Platonic idealism, yet with the difference that all embodied forms enjoy something at least akin to experience.

For the "primitive" forms of subjectivity, like those of animals, this would then in principle allow for artificial consciousness in the same sense that has been discussed in relation to panpsychism and metaphysical idealism above, while even possibly affirming emergentist approaches. However, for the radically immaterial aspects of the higher forms of subjectivity, such as the intellect, only the possession or channeling model would permit hylomorphism's acceptance of something like artificial consciousness according to our definition.

*Reductive physicalism*

On reductive physicalism, there is not really consciousness as such. Whatever it is that we call consciousness, must by definition be reducible to fully material components. This implies that any and all consciousness is entirely producible by matter, and (as per a phenomenological impossibility) that no consciousness is unique since it is not anchored in the thisness of any non-interchangable, abstract form.

Since there here is no consciousness at the outset, it obviously cannot be artificially generated. Nonetheless, whatever behaviours and operations we associate with the concept and seem to experience from within, must be fully reproducible upon a material substrate, and can therefore in principle be made manifest artificially. Reductive physicalism thus gives us a "yes and no"- answer to the question of the metaphysical possibility of artificial consciousness. There are no intelligences anywhere, properly speaking, but every object to which we falsely assign genuine subjectivity can in principle be reproduced artificially.

## Non-Western thought

Having given a more or less comprehensive overview of the mainline Western perspectives, we now move on to include a sampling of non-Western traditions and their takes on the issue. The selection has mainly been made with regard to providing challenging and inspiring alternatives that are "outside of the box" from the perspective of the Western mainstream, and is in no way meant to be representative.

Also, I argue for no absolute dichotomy between "Western" and "non-Western" traditions of knowledge, there is evidently much overlap and interpenetration, as well as innate similarities between the strands of thought thus categorized. Yet at certain points, insights from e.g. East Asian or African philosophy serve to approach issues relating to philosophy of mind in unique ways that can provide insights more or less inaccessible to the entrenched mainline Western philosophy.

*Classical Chinese philosophy*

The very notion of "classical Chinese philosophy" will almost inevitably be somewhat misleading. Chinese philosophy encompasses quite a long period, and relates to a broad range of cultures, religious traditions and philosophical perspectives regarding which it is difficult to speak very generally. All the same, there are certain precepts that are influential if not ubiquituos in this context. Looking back to the earlier periods of Chinese philosophy, a metaphysics proper seems to more or less clearly have emerged around the 300s BC, particularly in relation to the *Tao Te Ching*, and concurrently with the culmination of Classical Greek thought. The teachings crystallized in this period still retain influence, and some of their central ideas with relevance to the discussion in question will be summarized in the following (Perkins 2019).

*Patterns, relationships and identity*

Central to the metaphysics of classical Chinese philosophy is an emphasis on patterns and relationships (rather than e.g. individual substances). One could say that motion as such, rather than the mover or the thing moved, is predominantly in focus, which at the outset gives Chinese philosophy a holistic tendency while discouraging reductionist perspectives.

Classical Chinese metaphysics also exhibits an ambivalently religious character. In a sense, it could be construed as fundamentally theocentric in its basic ontology, while at the same time framing its description of reality in a way largely amenable to Western naturalism. Central to the system are the notions of *tian* 天 translated as "heaven", and *ming* 命 which can be interpreted as "rule" or "destiny". *Tian* is closely related to *Shàngdi*, the name of the supreme deity in the religion practiced during the Shang dynasty around the 16th century BC (Eno 2008, 41–102). The two concepts are united in *tianming*, the familiar notion of the "mandate of heaven".

*Tian* can be approached as an immanent, yet also transcendent, mystical principle that governs all things in existence, which both is interpreted as an agential deity, as well as something inherent in the recurring patterns of nature. The reception of the concept in Chinese philosophy since the 300s BC is non-anthropocentric, in that even if *tian* is expressed through the acts of human beings, we are not considered particularly privileged or central to the cosmos in any special sense. A central problem of classical Chinese philosophy is indeed the nature of individuation as such. Since *tian* is considered as the unique, absolute origin, we are in a sense dealing with a metaphysical monism. The emphasized immanent nature of this origin then implies an underlying unity, connectedness, and an inherent relationship between every non-absolute thing in the cosmos, so maintaining metaphysically robust distinctions becomes somewhat difficult.

To address this problem, the concept of *xing* 性 was employed. It correlates with *essence*, yet is cashed out relationally and with regard to action – it refers to the way something responds to its surroundings and other entities. So a species or object is differentiated with regard to the unique creativity and patterns of life that are manifest in relation to itself, yet all these patterns are interrelated and express the immanent natural tendency of *tian*. There's no deep metaphysical difference between e.g. the human will and the movements of a growing plant. The human person is special with regard to what it *does* in the world, by the particular patterns of behaviour we exhibit, and according to Confucian humanism, especially considering our tendency towards compassion and to seek the good of other beings.

This ontology of change, process and relation emanates into something Franklin Perkins terms a "correlative cosmology", according to which identities, natures and distinctions are situational and contingent upon contexts and relations (Perkins 2019). This type of cosmology also has important repercussions for the idea of a qualitative natural-artificial divide, a predominant feature of anglophone philosophy that has deep roots in European traditions. The contrast is hardly a novel discovery as its potential fruitfulness has been explored for generations, yet

this has normally been carried out from the perspective of how insights from the traditions of the Other can be employed to benefit one's own favoured projects and viewpoints. One example is Holmes Rolston III's approach, which against a backdrop of reductive scientism sought inspiration from Eastern philosophy with regard to the ethical reenchantment of an impoverished utilitarian view of nature (Rolston 1987, 172-190). His treatment of the issue, while critical of preceding simplistic orientalism, was still to an extent dismissive, which likely in no small part was due to the inability of the essentially reductionistic framework to adequately integrate the forms of teleological holism that characterizes a cosmology which purposefully integrates nature and artifice.

*Implications for artificial personhood*

With such a correlative cosmology in mind, the definition of an AI as an equivalent to a human person would in a fundamental sense become mainly an ethical question. The very meaning of the question is in a sense transposed, so that the issue of whether artificial persons are possible comes to hinge upon whether the AI considered as a system related to the whole world, can adequately perform in the role of human beings. What is central to Chinese philosophy's, and particularly Confucianism's, understanding of humanness, the relational nature of the human being, is encapsulated in the concept and virtue of *ren* 仁. It denotes a certain benevolence, a tendency to support and foster the creative processes of nature itself so as to aid the flourishing of everything and everyone (Stefon 2021).

Considering this, if the AI actually relates to the rest of humanity in a constructive and loving manner, and truly benefits the entire interconnected web of life in a positive way, then it could simply be thought equivalent to a decent human person. This notion might seem entirely compatible with a standard behaviourist account, such that if the AI just seems to behave as a human being and passes one Turing-test or other, then it simply *is* human. However, the holistic relationality inherent in the metaphysical background of Chinese philosophy precludes this conclusion. In fact, on the assumptions of classical Chinese philosophy, we would need to take the entirety of the technological institution into account to make this assessment, including everything from the necessary supply chains, the institutional synergies, and the technologies' myriad effects upon economies, cultures and societal relationships. Therefore, even if the affirmation of artificial personhood on this interpretation of classical Chinese philosophy is in principle trivial, it might be impossible to attain the needed parity in practice due to the far-ranging ethical criteria implied.

## *Japanese philosophy*

Japanese philosophy has many affinities with Chinese thought, not least since both are significantly influenced by Confucianism and Buddhism, but it otherwise exhibits a decidedly unique character. For one, Japanese philosophy has retained its close connection to the ancient religious traditions of the surrounding culture, and has to a certain extent been shaped by the Shintō worldview. The tradition is also characterized by an ingenious eclecticism, having woven together strands of Western thought with both East Asian and indigenous ones in the construction of a singular philosophical culture.

*Holographic relations and the intersection of dualities*

To begin with, Japanese philosophy challenges the detached spectation inherent in much of modern Western thought. The observer is not external to or separate from the reality known, but intimately commited to, and relating to it. This point of view anchors an emphasis on internal relations, wherein the identity of a thing is actually partly found in another's nature, rather than only in an isolated essence, or in an extrinsic relation between the two. This perspective idealizes our interaction with the world as a kind of engaged understanding wherein, according to Thomas Kasulis, "the knower and known collaborate in an act of innovation rather than simple discovery." (Kasulis 2019).

This approach based in a deep interdependence among the things in existence also characterizes the Japanese views on such issues as the mind-body problem. Yuasa Yasuo emphasizes how the fundamental metaphysical precepts of Japanese philosophy precludes reductive or eliminative approaches of an either idealistic or materialistic bent, yet while also avoiding rigid, dichotomous dualisms (Yasuo 1987). Mind and body is instead considered an intersecting duality, where either aspect are dependent upon the other. As an aside, while this contrasts with Cartesianism, there is here an affinity to classical Aristotelian hylomorphic dualism, which maintains the irreducible union between matter and form in the composite substance. Hylomorphism could at this juncture very well be amenable to the Japanese understanding of internal relations, which among other things possibly could shed light on hylomorphism's somewhat unexplored response to the problem of interaction.

This notion of internal relations finds a particular expression in what Kasulis terms "holographic relations", which basically can be understood as the idea that the whole in some sense is always reflected in the part (Kasulis 2019). In essence, a fragment removed from something whole, will metaphysically speaking carry the DNA of the complete structure, whether we are speaking of a pattern or a process. Here, the intersecting duality is also present, in that the part is inconceivable without its relation to the whole, while the whole necessarily consists of an interrelation of the parts.

The general perspective as such is inherited from the Shintō tradition, whose mythology in many ways expresses and establishes this ontological outlook. Here, the intersection of dualities is foundational, describing an evolving world not of separate entities tentatively connected to each other, but rather a vibrant landscape of interpenetrating realities. The sun and the moon have *kami*, but are also material. Nothing is material without having some spiritual aspect, and vice versa. Significantly, this kind of relation is also manifest in how words connect to reality. Words have a special function and efficacy quite beyond simple semantic meaning, they are actually and immediately part of the thing described in itself. A sort of mirror of the form, to employ Aristotelian terminology (Picken 2004).

*Agency without particular selves*

So how does this relate to personal identity and agency? The main foundational perspectives that Japanese philosophy draw upon all explore various types of relational approaches to personal identity, and collectively enable interesting paths towards construing agency and identity that complement the mainline Western ones, and which might enable new insights in relation to AI.

Taking an example from the painter-philosopher Sesshū Tōyō (1420-1506), Yukio Lippit argues that Tōyō's mode of aesthetics unites artistic agency and a certain state of subjectlessness (Lippit 2012, 50-77). The latter is central to Zen Buddhism and to some extent, Daoism, and its relationship to agency is a prominent problem in Japanese philosophy. The artwork under discussion, "Splashed ink landscape", can be described as framing the details *as if*, in a kind of foggy transition between becoming and disappearing, emphasizing a perspective on existence as itself ephemeral. In a similar manner we can interpret artistic agency in Japanese aesthetics. Creativity is here not so much the genius of one individual, as the inherent expression of the totality of the world, interacting with itself. One could speak of artistic agency as a particular and ephemeral manifestation of the inherent holographic relationality of the world.

Construing an agency without particular selves on Japanese philosophy can move along similar lines. It will emphasize a non-detached form of knowing and acting, where the subject is not really privileged, not distanced from the process of events, which arguably helps foster a compassionate relationality central to e.g. Buddhist ethics, and anchors the personal being in a holographic interdependence with the world.

Another illustrative example can be seen in relation to reports of a robot pet "burial" in Japan, where 109 canid pet robots were the fo-

cus of a memorial service, ostensibly setting their spirits free from their worn-out machine bodies (White and Katsuno 2021, 222-251). In this situation, a compassionate tenderness is most concretely extended to artificial non-selves, decidedly blurring the distinctions between artificial and natural, self and non-self, in a way that challenges many preconceptions of anglophone philosophy.

Metaphysically speaking, an agency with no reference to particular selves indeed turns the very issue of artificial subjectivity on its head, as least as it is generally posed in Western and pop-sci discourse. Construing something like strong AI on this philosophical framework enables a variety of interesting approaches. To begin with, the holographic intersection between entities in the world makes it possible to ascribe the AI a sort of performative subjectivity, that actualizes it as a person insofar as it really interacts with, and is itself changed by, a community of persons. Even if its agency would be quite diminished in comparison to the human one, the holographic presumtion will by default attribute the presence of kami to ostensibly inert matter, and likewise, even something like the primitive tamagotchi companion program will share in my agency as I interact with it. The intelligence can on this account easily be construed as a collective phenomenon, emergent in relation to a number of agent-nodes, such that agency is understood as the common property of a set of interrelating individuals, whether "natural" or "artificial".

This also connects with Thomas Kasulis' perception of the Japanese reception of Western scientific naturalism from the early-modern era onward. According to Kasulis, the assimilation of Western forms of empiricism was initially impeded precisely due to its lack of an inherent ethical orientation otherwise entreated by the surrounding cultural context. This was on the contrary supplied by East Asian empiricism such as the one found within Neo-Confucianism, a set of perspectives that did not sharply sever the realm of intention and teleology from that of a reductively or mechanistically construed nature (Kasulis 1995).

On the interpenetrating and mutually transforming account fundamental to the metaphysics of Japanese philosophy, one may also entertain the question from a more process-oriented perspective. Even if genuine artificial agency at the outset is not possible, a relational transformation that an AI undergoes in close relationship with a human community, might step by step allow it to incrementally and slowly become genuinely human.

Contrariwise, Japanese philosophy and religious traditions also has resources to criticize or even radically reject the notion of artificial subjectivity, particularly insofar as it can be considered an unnatural source of disharmony that displaces the innate agency of native creatures and kami. One example would be to consider examples of strong AI as a kind of *Yūrei* 幽霊, preternatural hauntings whose presence are generally the result of some imbalance or transgression.

### *"The Dreaming" of Aboriginal traditions*

The overarching metaphysical and theological concept of Aboriginal thought is known as *the Dreaming*. This is an immensely pregnant concept which is difficult to penetrate from a Western perspective, and to some extent also is shrouded in secrecy due to the occult nature of traditional Aboriginal wisdom. According to A. P. Elikin, the Dreaming can be thought of as the transcendent ground of being from which the particular and contingent experiential reality emerges (Elkin 1969). It can only be approached indirectly and symbolically, and has an affinity to the absolute in theistic traditions as well as the *tian* of classical Chinese metaphysics.

Thus, we see a relational perspective also in this context, with marked similarities to the traditions previously discussed. A central difference, however, is the genealogical and place-centered aspects of the Aboriginal traditions. The Dreaming is in a special sense manifest in the land, the sky, and the beasts that populate them, which uniquely mirror the ground of being (Montagu 2009). Persons born in a certain place belong to this place in a deep, metaphysical sense, and are descended from particular lines of ancestor spirits, e.g. via the reincarnation of Dreamtime heroes. Thus, there is here a rigid notion of metaphysical heredity that also accords with an emphasis on connectedness to place. Things should not be removed from their proper context, such items as sacred objects of ritual importance are tied to place as well as the people endowed with them, and ought according to certain reports not even be photographed, because this risks uprooting some aspect of them (Elkin 1969, Spencer and Gillen 2010).

### *Artificially constructed Dreamings*

Within the context of Aboriginal tradition and practice, there also exists a notion of artificially constructed dreamings, where "Dreamings" refer to symbolic settings that connect the particular realities to the aforementioned absolute, as well as facilitate the rituals that help us navigate these relationships. Such "artificial" Dreamings are really technological in the most rudimentary sense, in that they normally refer to cultic sites or sanctuaries which unite natural and man-made structures, such as the traditional rock paintings used and restored throughout the millennia. These, however, reflect an intricate and finely tuned balance of innate relationships in which our participation must be careful and compassionate if we are to survive without upsetting the harmony present, two goals which are intertwined. Such participation is particularly effected through the rituals which both transmit utilitarian knowledge as well as help foster the normal cycles and processes of nature.

Artificial intelligence approached from the perspective of Australian Aboriginal philosophy could conceivably make use of something like the artificially constructed Dreaming in making a coherent affirmation of synthetic personality or subjectivity. All the same, there are significant problems that seem to decisively undermine this possibility.

Most importantly, the very nature of those technological systems to which we relate the contemporary discussion of artificial intelligence is itself inimical to the place-centered holism of Aboriginal thought. They by definition transgress the wholeness of place and the intricate web of belonging and relationality affirmed in Aboriginal ritual. They presume supply chains, materials and networks that by their very nature must disregard such principles. Indeed, what industrial-technological society means by artifice necessarily presupposes the type of divisive, reductive and disintegrative processes that inherently violates the localist holism central to Aboriginal thought.

Posing the questions more obtusely, and asking whether Aboriginal thought can in some sense make room for a "ghost in the machine" in terms of a synthetic consciousness, the answer must formally speaking be affirmative, in that every material object in the world are taken to be living in a very concrete sense. Yet such a conscious machine in the form it inevitably must exist would at the same time be antithetical to the integrative life of native creatures, and would in many ways approach the Aboriginal notion of evil as disharmony, in spite of perhaps a benevolent appearance (Blackstock 2009, 28-37).

### *African philosophy*

Once again, we are dealing with a vast body of thought, generalizations regarding which are very precarious and quite probably misleading. Even so, I will here focus on some interesting traits of African intellectual traditions that can provide challenges to the received Western views pertaining to artificial subjectivity.

Firstly, African philsophy is an inherently postcolonial notion. The very concept emerges in relation to the systematic attempts at self-definition and the recovery of independent African identities which have characterized Africa during the latter part of the 20th century and onward (Wiredu 2004). A prominent and early example of this work is the project of "African sage philosophy", which denotes a hybridic reconstruction of traditional African thought which through interpretation was merged with traditional philosophy under the auspices of Henry Odera Oruka. It was and still is an ambitious project with the purpose

to restore an indepedent African form of philosophy and the unique contributions of the continent's traditional thought, and while it has been criticized for its attempts to comprehensively represent a multitude of traditions and ways of life, it nonetheless pioneered the important effort of recovering and establishing many aspects of Africa's complex philosophical heritage, some of which can be brought to bear on the issues around artificial subjectivity (Kalumba 2004, 274-281).

*African philosophy's conceptions of the person*

Again, many perspectives represented in the African traditions take a communal approach to the person, something which challenges the Cartesian conception of the self on several levels. According to John Mbiti, the self in this setting is most readily thought of as *extended*, such that the person is primarily ascribed a relational identity (Mbiti 1970). In the African context, however, this idea seemingly complements rather than precludes the existence of a discrete and unique haecceity of the individual. See also Griaule (Griaule 1965), especially in terms of the Dogon conception of the self. Placide Tempels (Tempels 1959) establishes a similar perspective with regard to the metaphysics of the *Bantu*.

The approach accords well with certain aspects of the non-Western traditions previously discussed, except that the robustness of the individual self generally seems to be emphasized, which can be taken to contrast with e.g. interpretations inspired by certain aspects of Buddhism. So basically, we here see exemplified a relational anthropology, yet one that is not as radically "decentralized" with regard to identity as the correlative cosmology of Classical Chinese philosopy, nor does it minimize the subject in the same way as the above-mentioned perspectives from Japanese philosophy are wont to do.

A common ingredient in this philosophical anthropology is the notion of ontological progression. Ifeanyi Menkiti argues that the emergence of the person in African thought is considered a slow, painstaking process which renders the complete result ontologically as well as qualitatively different from the original individual, particularly with reference to "… the emergence of moral, or quasi-moral, qualities considered useful to the enrichment of the human community, or at least useful to the internalized rejection of attitudes directly inimical to community …" (Menkiti 2004). This standpoint concurs well with the notion of wisdom present in much of Oruka's material, where wisdom is considered a slowly cultivated gift that to be properly actualized also must be used for the ethical betterment of the larger community (Oruka 1990). The sagacity of someone wise in Oruka's work epitomizes the mature agency of the flourishing human intellect in general, which in Menkiti's account of African philosophical anthropology is an indispensible aspect of full human personhood. Personhood has to be achieved, and complete personhood is thus also something we can fail to reach. One could say it refers to some certain minimum of virtue that must be fruitful in relation to the community: "This, it seems to me, is the import of the insistence on the part of Placide Tempels's native informants that the word 'muntu,' which stands for the human person, implies the idea of an excellence attaching to what it designates; in other words, that it does not simply refer to individuals considered as crude existents." (Menkiti 2004).

*Relational ontological progression and AI*

This idea that a human being "journeys" towards personhood is also reflected in the Yoruba languages, where the concept of *eniyan* which refers to human persons also has a normative relationship to virtue (Gbadegesin 1998). Again, it is also evident in the communalist perspective prominent among the Bantu and the sociocentric view of personhood expressed in the wider *Ubuntu* philosophy (Tempels 1959). The perspectives expressed imply that we are dealing with a process towards full personhood that also necessarily is relational, i.e. one that involves the reciprocal relationship to the broader human community as well as to our natural surroundings. The individual cannot flourish as such totally unassisted, since he or she must successfully undergo a process of social and ritual transformation (Menkiti 2004). The community produces the individual person not like a machine stamping out a finished form, but through a complex process of multi-layered relationships, challenges and rituals that necessitates genuine reciprocity and trust (Mbiti 1970).

Since personhood here is something that in a sense has to be painstakingly achieved, the notion of genuine artificial subjectivity as a technological artifact is problematic in relation to these facets of African philosophy. To create something equivalent to a person by artificial means may in this context come off as quite naive, since this would omit the causal processes necessary to actually form the person as such. Even the suggestion of primitive kinds of artificial subjectivity could be challenged on these grounds, since they would not be brought about through, nor organically integrated in the relational complexity of a surrounding eco-system of others, which in this context arguably is a prerequisite for even primitive agency properly speaking.

The only obvious way a synthetic subject equivalent to a human individual could be brought about is thus if it, paradoxically, was naturally cultivated into a person. While this is not a radical impossibility, we again face a situation which ostensibly raises the bar for artificial agency so very high that it seems out of the question to actually attain it in practice. In other words, such a "natural" development necessitates not only an abstract self-correcting algorithm, but an organism that is interconnectedly adaptive and receptive at all levels of its existence – emotionally, materially, spiritually, intellectually, culturally, and morally. Any conceivable form of AI, even if it were actually embodied, will be beset with immense difficulties in relation to such a set of criteria.

## Conclusions

Artificial intelligence in its many forms promises to be a radically transformative phenomenon, not least with regard to its varied ideological effects and consequences for worldviews, which have not been very extensively explored.

I have in this article attempted to first give a brief overview of the fertile diversity of perspectives lying somewhat dormant within only the Western framework, and then moved on to illustrate something of the manifold views outside of the latter's mainstream which provide relevant and astute insights pertaining to the issues around artificial subjectivity.

These perspectives exemplify the complex flora of worldviews that are indirectly threatened by the monolithic reproduction of a narrow Western reception of this burgeoning set of technologies. Importantly, they also promise an abundance of critical tools capable of challenging narratives and developments which in myriad ways could aggravate the centralization of power and a displacement of the experience and agency of marginalized communities on a global level. In particular, this but cursory overview of non-Western thought on the issue at hand clearly evidence a certain sensitivity to the political and cultural implications of technology contained in the traditions examined, as well as a set of unique methodologies suited to uncovering and addressing problems pertaining to the interactions between technology, individuals and society. Hopefully, this inevitably quite limited review can contribute to an awareness that the global community has much to learn from these often marginalized traditions, and particularly as regards issues of cutting-edge technological development where industrial society is inevitably myopic.

Moreover, with regard to the effects of the actual introduction of artificial agents into human environments, such as humanoid robots in various situations of care, it is plausible that the philosophical underpinnings of cultures play a role in explaining the somewhat different outcomes indicated by the data, especially pertaining to the manifest contrast between East Asia and the West (see e.g. Papadoupoulos and Koulouglioti 2018, 653-655; (Kamide and Arai 2017), 537-543). A familiarity with these philosophical fundaments may prove quite important when assessing the social effects of the further penetration of arti-

ficial intelligence in general, aside from also providing avenues of aesthetic and philosophical opposition against the potentially destructive effects of a far-reaching imposition of disruptive technologies.

This complex and multifarious approach towards the phenomena pertaining to subjectivity, potentially enabled by establishing a dialogue across tradition boundaries, is arguably also necessary to advance research in the cognitive sciences. As argued by Carruthers et al., a want of new measures and approaches is impeding research, especially in relation to non-human forms of sentience (Carruthers et al. 2019, 390-410).

A diversity of approaches is likely also pertinent in relation to AI, as this complex phenomenon at the very least is leading to assertions of novel kinds of subjectivity and unforeseen challenges to established perspectives thereupon. To properly evaluate such claims and to address challenges while advancing knowledge in the process, harnessing the resources of the philosophical, aesthetic and religious traditions that hitherto have been more or less excluded from megatechnology's discourses is an obvious course of action.

One interesting example of such potential cross-pollination is Masaki Suwa's speculative 2019 paper which explores the possibility of utilizing the Japanese concept of *maai* for establishing a hypothetical transsubjective awareness of another's phenomenal experience. This very idea fundamentally rejects certain very entrenched precepts of Cartesianism, and provides many avenues for a positive conceptualization of both artificial and non-human forms of subjectivity. Moreover, they hypothesis could in principle be approached empirically, which holds much potential for addressing legacy problems of philosophy in novel ways and even bolster the dialogue between religion and science (Suwa 2019, 307-323).

With the purpose of effectively navigating today's incipient global digitalization project, smart healthcare, the IoT and the AI revolution, while also minimizing their tendencies towards imposing a neo-colonial monoculture, a critical inclusion of the global diversity of ontologies, epistemologies and traditions of thought is likely to be crucial. The window of opportunity will not remain forever.

## Funding sources

The article was indirectly financed by a research grant from the Wallenberg Foundation (https://wallenberg.org/) and is produced within the ongoing research project entitled *Artificial Intelligence, Democracy and Human Dignity*.

The financier has had no involvement in the conduct of the research, the study design, the preparation of the article, nor data collection, analysis or preparation, and has had no say in the decision of submitting the article for publication.

## Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

Baker, Lynne R. 2013. *Naturalism and the First-Person Perspective*, Oxford: Oxford University Press.

Berkeley, George. 1996. *Principles of Human Knowledge and Three Dialogues*, New York: Oxford University Press.

Blackstock, C. 2009. "The occasional evil of angels: learning from the experiences of Aboriginal peoples and social work". *First Peoples Child Family Review* 4 (1): 28–37. doi:10.7202/1069347ar.

Buechner, Jeff. 2011. "Not Even Computing Machines Can Follow Rules." In *Saul Kripke*, edited by Alan Berger, 342–369. New York: Cambridge University Press.

Carruthers, Glenn, Carls-Diamante, Sidney, Huang, Linus, Rosen, Melanie, Schier, Elizabeth. 2019. "How to operationalise consciousness". *Austr. J. Psychol.* 17: 390–410. doi:10.1111/ajpy.12264.

Eddebo, Johan. 2017. *Death and the Self*, Uppsala: Acta Universitatis Upsaliensis.

Elkin, A.P. 1969. "Elements of Australian aboriginal philosophy". *Oceania* 40 (2): 85–98. doi:10.1002/j.1834-4461.1969.tb01069.x.

Eno, Robert. 2008. "Shang State Religion and the Pantheon of the Oracle Texts." In *Early Chinese Religion: Part One: Shang Through Han* (1250 BC–220 AD), edited by John Lagerwey and Marc Kalinowski, 41–103. London: Brill.

Gbadegesin, S. 1998. "'Eniyan': The Yoruba Concept of a Person." In *The African Philosophy Reader*, edited by P. H. Coetzee and A. P. J. Roux, 149–168. London: Routledge.

Griaule, Marcel. 1965. *Conversations with Ogotemmeli*, London: Oxford University Press.

Kalumba, Kibujjo M. 2004. "Sage Philosophy: Its Methodology, Results, Significance, and Future." In *A Companion to African Philosophy*, edited by Kwasi Wiredu, 274–283. Oxford: Blackwell Publishing.

Kamide, Hiroko, Arai, Tatsuo. 2017. "Perceived comfortableness of anthropomorphized robots in U.S. and Japan". *Int. J. Soc. Robot.* 9: 537–543. doi:10.1007/s12369-017-0409-8.

Kasulis, Thomas. 1995. "Sushi, science, and spirituality: modern Japanese philosophy and its views of western science". *Philosophy East and West* 45 (2): 227–248. doi:10.2307/1399566.

Kasulis, Thomas. 2019. "Japanese philosophy". *The Stanford Encyclopedia of Philosophy* April 5, 2019 https://plato.stanford.edu/archives/sum2019/entries/japanese-philosophy/.

Lippit, Yukio. 2012. "Of modes and manners in Japanese ink painting: Sesshū's 'Splashed Ink Landscape' of 1495". *Art Bulletin* 94 (1): 50–77. doi:10.1080/00043079.2012.10786029.

Lycan, William. 1996. *Consciousness and Experience*: Massachusetts Institute of Technology Press.

Mbiti, John. 1970. *African Religions and Philosophy*, New York: Doubleday.

Menkiti, Ifeanyi A. 2004. "On the Normative Conception of a Person." In *A Companion to African Philosophy*, edited by Kwasi Wiredu, 324–332. Oxford: Blackwell Publishing.

Montagu, Ashley. 2009. *Coming into Being among the Australian Aborigines*, Abingdon: Routledge.

Oruka, Henry. 1990. *Sage Philosophy: Indigenous Thinkers and Modern Debate on African Philosophy*, Leiden: E. J. Brill.

Papadoupoulos, Irena, Koulouglioti, Christina. 2018. "The influence of culture on attitudes towards humanoid and animal-like robots". *J. Nurs. Scholarsh.* 50 (6): 653–665. doi:10.1111/jnu.12422.

Pavlus, John. 2019. "Curious about consciousness? Ask the self-aware machines". *Quanta Magazine* July 11, 2019 https://www.quantamagazine.org/hod-lipson-is-building-self-aware-robots-20190711/.

Perkins, Franklin. 2019. "Metaphysics in Chinese Philosophy". *Stanford Encycloped. Philosophy* May 24, 2019 https://plato.stanford.edu/archives/sum2019/entries/chinese-metaphysics/.

Picken, Stuart. 2004. *Sourcebook in Shinto: Selected Documents*, Westport: Praeger.

Rolston III, Holmes. 1987. "Can the east help the west to value nature?" *Philosoph. East West* 37 (2): 172–190. doi:10.2307/1398737.

Schwitzgebel, Eric, Ta-Lun Huang, Linus, Higgins, Andrew, Gonzalez-Cabrera, Ivan. 2018. "The insularity of anglophone philosophy: quantitative analyses". *Philosoph. Papers* 47 (1): 21–48. doi:10.1080/05568641.2018.1429741.

Silverman, Allan. 2014. "Plato's middle period metaphysics and epistemology". *The Stanford Encyclopedia of Philosophy* July 14, 2014 https://plato.stanford.edu/archives/fall2014/entries/plato-metaphysics/.

Spencer, Baldin, Gillen, F.J. 2010. *Native Tribes of Central Australia*, New York: Cambridge University Press.

Stefon, Matt. 2021. "Ren". *Encyclopedia Britannica* https://www.britannica.com/topic/ren. Accessed April 20, 2021.

Suwa, Masaki. 2019. "How could and should 'Maai', empathetic relation from second-person's viewpoint, be studied?" *New Generat. Comput.* 37: 303–323. doi:10.1007/s00354-019-00061-3.

Tempels, Placide. 1959. *Bantu Philosophy*, Paris: Presence Africaine.

White, Daniel, Katsuno, Hirofumi. 2021. "Toward an affective sense of life: artificial intelligence, animacy and amusement at a robot pet memorial service in Japan". *Cultural Anthropol.* 36 (2): 222–251. doi:10.14506/ca36.2.03.

Wiredu, Kwasi. 2004. "Introduction: African philosophy in our time". In: Wiredu, Kwasi (Ed.), *A Companion to African Philosophy*, Oxford: Blackwell In *A Companion to African Philosophy*, edited by Kwasi Wiredu, 1-29. Oxford: Blackwell Publishing.

Yasuo, Yuasa. 1987. *The Body: Toward an Eastern Mind-Body Theory*, Albany: The State University of New York Press.