

The dream of transcending the human through the digital matrix: A relational critique

PIERPAOLO DONATI*

University of Bologna, Bologna, Italy

pierpaolo.donati@unibo.it

ORCID: 0000-0002-9794-132X

Abstract. The advent of the digital era brings with it the dream of ‘transcending the human’ through the most sophisticated AI / robot technologies. The Author argues that the concept and practices of ‘transcendence’ are deeply ambiguous, since on the one hand they simply aim to overcome the weaknesses, limits and fragility of the human, while on the other hand they modify the human by selecting its specific qualities and its causal properties in a way to generate beings ‘other than human.’ Post / trans-humanist ideologies supporting a radical transformative change, instead of strengthening the human, lead to its dehumanization because they imply the denial of the characteristics that are uniquely human and / or those that constitute human nature. To understand when new digital technologies enhance the human (keeping it within its natural boundaries) or distort it, it is necessary to analyze and evaluate social processes on the basis of a relational vision that shows us which human relationships emerge from the use of technologies.

Keywords: human transcendence; digital technologies; personhood; dehumanization; relational sociology; relational goods.

* *Alma Mater* Professor of Sociology, Dept. of Social and Political Sciences, University of Bologna.

1. The issue: “transcending the human”

Since prehistoric times, humans have lived in an ontological condition of profound ambivalence: on the one hand they experience their own limitations, and, on the other, since they do not accept this condition, they want to go beyond their own nature. They feel they are made for the infinite, but every day they have to struggle in order not to fall back into the limits of the finite. They are the most unhappy beings, because, being weak and fragile, they dream of living without limits. The advent of the digital age seems to provide a new horizon of opportunities to take a decisive step towards the possibility of “transcending the human.”

What does it mean to “transcend” the human? Certainly it does not simply mean the possibility of using instrumental artifacts that can exempt people from certain tasks and functions, giving them more freedom (according to Gehlen’s 1974 *Entlastung* theory). It does not mean that people can have new machines available to help them in doing their work. It means much more. It means going “beyond the human.”

The quest to transcend the human by what I call “digital technological matrix” (Donati 2020) does not represent the creation of a “more human” being, but modifies the human by selecting its specific qualities and its causal properties in a way to generate “more than human” beings, which means “other than human.” From the cultural point of view, I define the digital technological matrix as the globalised symbolic code from which digital artefacts are created in order to help or substitute human agency by mediating interhuman relations or by making them superfluous. From the structural and practical point of view, the digital matrix is the complex of all digital technologies based on scientific knowledge and AI engineering (Donati 2021).

To put it bluntly, I argue that the idea of transcending the human through the digital technological matrix as envisaged by posthuman, transhumanist and cyborg ideologies that support a radical transformative change, leads to an increasing dehumanization.

Although the concept of dehumanization lacks a systematic theoretical basis, I maintain that, in general, it refers to meanings that involve the

denial of the *characteristics that are uniquely human* and / or *those that constitute human nature*. Most of what we call post-human theories and praxis entail cognitive underpinnings of animalistic and / or mechanistic kinds that do not represent forms of human transcendence, but instead of dehumanization. In the face of the trans-human era, we need new criteria in order to evaluate what humanizes the human and what de-humanizes it.

In my view, the human person does not transcend herself in her single action¹ because, even if augmented by new technologies, in her single action the person simply remains herself. The person transcends herself, that is, she goes beyond herself, when she comes out of herself, that is, she transcends herself *in the inter-human social relationship*.

From my point of view, the transcendence of the human consists in a relational capacity and not in the capacity of an individual entity like a robot (Donati 2020). For this reason, I cannot share the perspective of those scholars who argue that we find ourselves in front of a person (not necessarily endowed with a human body) if the mind of such an entity (for example a robot) is capable of activating individual reflexivity in action. This thesis is supported for example by Archer (2019a), who admits the possibility of creating robots endowed with the first-person perspective. She believes that robots can be friends of the human beings and, as such, should be attributed human and citizenship rights. To me, the arguments put forward by Baker (2002) and Archer (2019b) reflect a severely deficient anthropology in understanding what it means to be able to transcend the human.

From my point of view, it is *in* the interpersonal relation, *with* it and *through* it, that we can understand what is *indefungibly* human, both in the single person and in social forms. The human, as a reality *in actu*, lies in the intersubjective relation because the latter has qualities and causal properties that are not replaceable by any other entity, be it artificial or belonging to the natural physical world. When someone speaks of the virtues of the fidelity of dogs or the humility of chimpanzees, and more generally of certain animals

¹ By 'single action' I mean the Parsonian 'unit act' (see Parsons 1937).

as ethically superior to the human being, this is to talk about qualities and causal properties that reside in one's mind, not in reality, because human virtues are characterized by moral qualities that non-human animals cannot in any way have. The same holds true for smart robots, androids or other human-like machines. When a child talks about the robot as his 'best friend,' he is talking about a relational experience that does not have the qualities and causal properties of an inter-human friendship relationship, but it is only a series of exchanges that can be useful and satisfying for him.

2. Rethinking human evolution

In my view, we have to face two grand scenarios of human evolution, underlying the present cultural dynamics in the globalized world, that implicitly contain two prevailing models of understanding human transcendence. On one side, a (Western) scenario based upon the expansion of science combined with technology, positivistically geared to generating new being with superhuman abilities. On the other side, the (Eastern) idea that world evolution consists in a continuous bringing into existence of what does not exist, i.e., a process in which human transcendence means the construction of a machine of Nothing. As Jones (2013, p. 418) reminds us, Western traditions (with exceptions such as Spinoza) view the world as an array of individual objects that are bound in a web of causal ties, whereas Eastern traditions view it as a unity in which the appearances of plurality and diversity are no more than ripples on the surface of an oceanic continuum. Applying this Eastern worldview (called 'Buddha-nature') to the human-machine interaction, depending on how you look at it, I could be regarded as managing the automobile, or it could be regarded as managing me. To control, in effect, is to be controlled. Thus, human beings and machines are fused together in an interlocking entity.

(a) The first scenario can be exemplified by those contributions to the volume *The Onlife Manifesto; Being Human in a Hyperconnected Era*, edited by Luciano Floridi (2015), which assume that human nature is equal to a *bios* equipped with psychism and that in the future the human

will consist of these dimensions enhanced by technology. The distinctions between human beings and other living beings (animals and plants) will be gradually reduced. Human sociability will be transformed accordingly. In this perspective, the human being is considered as *nature's beautiful glitch*, i.e., a beautiful error of natural evolution that has succeeded in surviving and now is evolving into an info-sphere which becomes anthropo-*eccentric* rather than anthropocentric. The basic tenet is that '*more tech makes us more human*'. Such a prospect runs the risk of losing a critical view of the different forms of enhancement, that is, of losing the distinction between those technologies that allow the human person to maintain her identity and those that produce deterioration, such as to break or violate that identity making the human person a slave of the same technologies.

It seems to me that opposing Floridi's view on the basis of a wholly individualistic first-person perspective is not enough, because the latter does not recognize the consistency, needs and uniqueness of the relation between mind and body in each person. If one follows Lynne Rudder Baker (2007, p. 205), one could think that the first-person perspective is perfectly compatible with the positivistic scenario depicted by *The Onlife Manifesto*, once we agree with Baker that technology can change the human body without affecting the person's identity, since the person can still think of herself as herself.

If parts of my human body were replaced by synthetic parts until the body that constitutes me was no longer a human animal, then, as long as my first-person perspective remained intact, I would continue to exist and I would continue to be a person. But if nothing had my first-person perspective, then there would be no me. (Baker 2002, p. 371)

My account of the first-person perspective has some naturalistic and some non-naturalistic aspects. It is naturalistic in that it does not appeal to immaterial souls. The first-person perspective may well have evolved by means of natural selection; we human persons, with our first-person perspectives, are as much a part of the natural world as were dinosaurs. I have no doubt that there's something going on in my brain that makes it possible for me to have I* thoughts [in Baker's language, the asterisk means thoughts of myself on

myself] and I have no doubt that our capacity to have I* thoughts is a product of natural selection. (Baker 2007, p. 206)

I agree with Baker that there cannot be an adequate ontology—an inventory of what really exists — that includes no first-person subjects of experience. But to agree on this perspective does not mean accepting the total splitting of body and mind in a person, because it would mean underestimating the influence of social relationships in the constitution of personal identity. It would mean a return to the Cartesian '*cogito, ergo sum*'. That is, I exist only because I think, whereas my body is held to be a '*res extensa*' (extended thing) possessing the substance of wax. According to this argument, one could never similarly demonstrate the existence of the 'other'. Because of this, humans can know the self and its thoughts, but cannot actually know anything of anything that is not the self.

Human body, mind, and agency cannot be completely separated. Either they have some integration, or the human is no longer such and the action is no longer human. The unity of this difference is given by the energy (*energeia*) proper to their relationship. Splitting body and mind is an abstract operation that can serve to justify a individualistic first-person perspective, defined as “our inwardness—our abilities not just to think, but to think about our thoughts; to see ourselves and each other as subjects; to have rich inner lives.” Baker 2002, p. 370). This is a naïve perspective. It does not see how the relationships between people influence their inner life, and does not provide any meaningful account of the real experiences of 'normal' people (not humanoids or robots) in connecting their body and mind in a hyper-technological society. The 'inner' abilities Baker is referring to are strictly dependent on the new communicative environment as well as the possibility of keeping a sense of identity in the connections between body and mind. Baker's first-person perspective seems to be a strategy that proposes to leave the body to technologies and preserve (indeed enrich) the 'reflective mind,' as if the same technologies did not have a powerful influence on personhood through body modifications. In the end, what she calls the “constitution view” of the human person is *a-relational*, (Baker 2016)

in a double way: because she splits what is human and what is person, and because she ignores the role of social relations in constituting the way in which the first-person perspective can actually operate.

On the Constitution View, something is a *person* in virtue of having a first-person perspective (or a narrowly defined capacity for one), and something is a *human* person in virtue of being a person constituted by a human animal (or body). Human persons are material beings, part of the natural order. As I develop the idea of constitution, this view of human persons has the consequence that although I am both a person and an animal, I am most fundamentally a person. Hence, my persistence conditions are the persistence conditions of a person (sameness of first-person perspective), not the persistence conditions of an animal (sameness of biological organism). (Baker 2002, p. 370)

(b) The second scenario can be illustrated by quoting an evocative paper by Bilimoria (2012), in which he comments on the Vedic texts, where it is written: “In the beginning, there was neither Non-Being (asat) nor Being (sat)” (RgVeda X.129.1; it is called ‘hymn to the creation’). Bilimoria asks: “what was there, bottomless deep?.” Eventually he responds positively. The argument is that “Radical Nothingness ought not to be feared: it may have therapeutic value in the hallowed Wittgenstenian-Yogacārin sense (an antidote for excessive linguistic conceptualism); it may prove to be a timely yet fairest watchdog on the excesses of Technoscience culture (to which Heidegger similarly drew attention), as well as on fundamentalisms of all varieties the return of religion to the public sphere (despite the Enlightenment’s decree), secularism re-seeking the sacred enchantments, and on other attachments to Absolutes.” (Bilimoria, 2012: 528). Therefore – so he concludes, “let there always be non-being so we may see their subtlety, and let there always be being so we may see their outcome. The two are the same. But after they are produced, they have different names. They both may be called deep and profound... Deeper and more profound...” (quotation from *Tao-te Ching* by Lao-tzu).

Ultimately, this scenario leads to see the evolution of the whole of creation, and of human evolution within it, as a ‘machine of Nothingness,’

in the sense of bringing into existence what does not exist originally. In the end, Bilimoria describes his perspective as a “frustrated attempt to complete fabricating a machine that creates Nothingness, or work it up at least as the transcendental concept of which no smaller can be conceived” (Bilimoria 2012, p. 529).

In its naturalistic version it could be, for example, bringing into existence and socially institutionalizing a sexual role which is neither male nor female, as in the case of *fa’afafine* in Samoa culture. In Japanese, Chinese and Indian cultures it can mean to give real existence to certain fictional beings. In this scenario, there is no reason to distinguish between Human and Non-Human, or to think of transcending the human, simply because, to put it bluntly, the religious matrix of these cultures is a sort of naturalistic pantheism. Their transcendental matrix is the same Nature. I use the term ‘transcendental matrix’ in the same sense that can be found in Weber’s analysis of the religious roots of world cultures (Weber 1993).

The general notion of ‘matrix’ here refers to a cultural code (symbols and norms for their use) that orients or steers the way in which agents/actors intersect their intents, the situational context in which they find themselves, the meanings they use and the generated effects.

... intent, context and effect are better considered as intersecting concepts that cannot be compartmentalized into discrete categories. It is not about privileging intent, context or effect, but recognizing their interdependency. In order to move away from the notion of discrete variables that have their own unique influences, we call this inter-relationship the intent-context-effect matrix. (McCormack et al 2016, p. 760)

While the term matrix is generally used in the sense of ‘matrix of society,’ here I would like to underline its transcendent, metaphysical dimensions compared to a constructionist view of society.

In short, I have chosen to mention these two visions because, although distant and even contradictory in some respects, they seem to merge in globalizing society. What emerges from their mix is a culture that views human transcendence as an endless evolution that is supposed to supersede

human deficiencies by mixing natural and artificial elements of all kinds. Western rationalism is mixed with Eastern naturalism within a complex 'nature + technology' in which transcendence is a matter of immanence. The result is a vision of human transcendence as a process of continuous up-grading of life forms which resembles a process of incessant creation of new singularities (see the theory of technological singularity). From my point of view, however opposite, these two scenarios seem to be connected by a subtle and hidden thread, that of a kind of *vitalism* mixed with some kind of *existential relativism*. What is now called '*post-truth politics*', is, in fact, a form of *cultural post-rational and post-ethical agnosticism*, since in the digital media 'anything goes', which means that communications do not discriminate between truth and non-truth, and rationality is substituted by emotionality.

The two scenarios just mentioned place human transcendence, respectively, in the total immanence of technological evolution and in an immanent process of creation that makes exist what is not. It is noteworthy that this strategy was adopted by Niklas Luhmann (1995) to build his overtly anti-humanistic sociological theory (Donati 2011).

I would like to present here a third scenario that introduces another idea of human transcendence. It conceives transcendence as an emerging relation between what exists (immanent reality) and what can be (transcendental reality). What I call the 'relational order' is precisely the transcendental, generative mechanism linking immanent and transcendent realities.

Bhaskar's basic argument, according to which the 'true' objects of social sciences are the *real* (although typically unobserved) generative mechanisms underlying phenomena (and not what can be observed and measured empirically), is transcendental in nature, i.e. it implies a transcendental logic array (Bhaskar 1989). Later on he talked of a meta-reality. Those who deny this possibility end up by erasing the possibility of any sociological humanism. This is what Breslau (Latour's follower) claims: "... human agency depends on the agency of things, and the objectivity of the social world depends on the objectivity of the natural world-as we know it, not an unknowable transcendent natural world." (Breslau 2000, p. 305).

I argue that at the beginning of any creation (not only of the original creation) there is not a bottomless depth, but a Being ever able to transcend itself by relating himself to Non-Being. It is in this process of transcendence (emergence) that being and non-being are related, while no one of them can exist *per se* in absolute isolation. The very nature of creation would then consist in a relation that brings into existence what does not exist through a process of emergence. Is it possible to think that what is human can transcend itself in this ‘relational nature’ of creation? The possibility of arguing in favour of a positive answer to this question lies in the social ontology of what we mean by ‘relation’ and ‘relationality.’

3. The distinction human/non-human as a transcendental relation, and its enigma

To speak of transcendence implies a social ontology of reality. In turn, such an ontology refers to a transcendental reality, and therefore it calls for the existence of a transcendental cultural matrix to which the ontology is closely connected. When I claim that this matrix is ‘theological’ I do not refer to the dogmatic beliefs of a specific religious faith. I mean the *symbolic code* underlying every great culture or civilization concerned with ultimate realities (Donati 2019a).

It is a matter of fact that any science does refer to such a kind of cultural matrix, although very often in an unwitting or unspoken way. All sociologies have a symbolic matrix that depends on a ‘mother-matrix’ where ultimate realities are placed. According to Anders (2016, sections 22–26), the fact that not only our experiences, but even our needs, are molded, represents the maximum service of the matrix. This is because the matrices mold not just us, but the world itself. The artificial models of the ‘world’ not only mold us and our image of the world, but the real world itself. This molding process has a boomerang effect: that the lie really lies to itself. In short: *the real is transformed into the copy of its images*. Many scholars will probably disagree on this claim, but nevertheless they cannot deny that any scientific theory possesses a meta-theory through which it ‘sees’ the world. The founding

fathers of sociology, including Comte, Marx, Durkheim, Weber, provide different examples.

For clarity's sake, let me make explicit what I mean by the term 'transcendence.' To me, it means: (i) a *meta-level of reality* that we need to reach in order to get a deeper explanation of what happens in reality or to imagine a reality beyond what exists, by relating to a further 'layer' beyond what already exists in an open horizon (matrix) of up-grading orders of realities; (ii) the *necessity of this relation*, as distinct from its contingencies; (iii) whereas 'necessity' is a matter of recognition of what is *latent* (hidden, invisible, unmeasurable) as an 'ultimate reality' bordering on religion (i.e. what is sacred, and, as such, conferring a *dignity* to every being, in the sense of recognizing that every existing being deserves the *respect* that is due to it according to its own nature). The recognition of the dignity of every entity (in Latin the word *dignus* means 'a thing or a person deserving respect for its qualities') marks the boundary between what I call '*immanent transcendence*' (as a product of society) and '*transcendent transcendence*' (as a reality which is not the product of society, because it is pre- and meta-social, it is a meta-reality).

The concept of human dignity can become significant only if one admits the existence of these two orders of reality (what transcends factual reality for emergence from empiricism, and transcendence that does not depend on facts because it is meta-empirical; Bhaskar 2012), and only if one is able to see and manage the intrinsic relationships of these two orders. In the absence of these conditions, proposals regarding a good society sound in vain or they are reduced to abstract moral and legal notions. This is the case, for instance, with three reports issued in October 2016 by the White House, the European Parliament, and the UK House of Commons outlining their visions concerning how to prepare society for the widespread use of AI. After assessing the merits of these reports, Cath et al. (2017) complain that they do not provide any foresight for describing the future that, as a society, we would like to see. Cath et al recommend a two-pronged approach in order to steer the process of developing the 'good AI society.' According to them, on the one hand, policies should ensure that AI is steered fully towards

promoting the *public good*; on the other hand, projects could fruitfully rely on the concept of *human dignity* as the lens through which to understand and design what a good AI society may look like. But these authors have to admit that both notions of ‘public good’ and ‘human dignity’ are latent realities, irreducible to their immanent definitions in legal or otherwise terms. These realities belong to the relations between the domain of immanent transcendence and meta-reality, i.e. to the latency of society.

This latent reality should not be understood as a *Deus ex machina*. Such a mechanical God is *not* necessary for the final resolution of human enigmas as in Greek tragedy. The resolution lies in adopting a *relational theological matrix* that can make us understand the reasons why we need to resort to some kind of transcendence in order to explain the dynamics of social reality without mechanical or deterministic artifices. I define a symbolic code (or semantics) as a set of symbols and the rules to use them in looking at the world and interpreting phenomena, facts and events. When the symbols refer to ‘ultimate realities’ (ἔσχατος, *éskhatos*) and the rules follow a logic of first principles, a symbolic code takes on the form of a ‘theological matrix’. Why I call it so? Because, in this case, the symbolic code is a reflection of a theology (*theos+logos*), i.e. a discourse or inquiry on the divine truth. In short, the symbolic code derives from the way we ‘semantize’ God, where people find His presence in terms of those ultimate realities that explain what happens in the world.

The question we are faced with is: what are the relationships between the human and transcendental realities?

We are confronted with a crucial, terribly difficult issue, which is the following: whether or not, to support determinate human-social relations (‘horizontal relationships’, so to speak), a transcendental matrix is necessary (that is, resorting to ‘vertical’ relations among the stratified layers of reality). If not, why? If yes, which characteristics should it have? More generally, the question is whether or not the *social* ontology of the relations of which we speak (like, for example, love, as an interpersonal relation and not as a subjective sentiment or passion) requires a metaphysics rooted in a theology understood not as a discipline in itself and for itself (based on

Revelation), but, in sociological terms, purely as a symbolic cultural matrix. In this paper I argue that the source of the *meaning of the human- social relation* lies in its transcendental matrix, although the social relation is generated (or is not generated) due to autonomous causes.

For example, everyone agrees that love is essential in the life of a couple or a family. But love must be qualified, it cannot be a relationship whatever: it must be love in a due order and with a ‘deep’ meaning. In such contexts, as an ultimate concern, love needs to be referred to a transcendent symbolic matrix from which it takes its meaning in order to be pursued with appropriate practices. From a Simmelian point of view, which I share, social life has a transcendent character (“the transcendent character of life”) that permeates even the use of money, given the fact that, in Western culture, mundane monetary interactions and exchanges are related to key religious forms of experience, like faith, unity, and individuality (Silver and O’Neill 2014).

In this contribution, I would like to say how relational sociology seeks to respond to these basic questions on which the possibility of linking human (social) relations and transcendental (supernatural) relations depends. This possibility, in its turn, depends on the fact of being able to see and manage the enigma of the relation, the enigma that lies in relationality as such (Donati 2019b).

In summary, the problem is the following: whether or not it is necessary—if not, why, and if so, what is it—to resort to a *symbolic matrix* that allows us to face the enigma of the relation in such a way that it is possible to see how and why human relations and transcendental relations are *at the same time ontologically connected* to each other (not by similitude or metaphor) and *distinct*. This is possible to the extent that the said symbolic matrix assumes that the relational is a transcendental or genetic principle (Debaise 2012).

The inherent defect of traditional humanism, based upon naïve realism, is that of understanding social relations with others and with the world simply as a “manifestation” of the qualities inherent in the human person and her inner life. Of course, good actions are the energy to break out of ourselves and

be prepared to undertake generous tasks which will be of benefit to all. However, to undertake an action on the basis of internal impulses does not mean that a certain result actually follows. It poses the question: are human-social relations simply an expression of inner experiences, of self-consciousness, in one word of the internal conversations of persons? Or are social virtues simply inherent in human nature? In my view, in answering these questions (which are the enigmas that social reality, as an oracle, places before us), we must face unexpressed challenges whose solution lays the path that can lead to a possible neo-humanism open to transcendence.

As is well known, the classical (Aristotelian) conception affirms that virtue does not depend on consciousness (this corresponds rather to the modern, Cartesian or even Kantian, idea), but rather on good action, that is, in accordance with one's own nature; according to this view, virtue ultimately depends on nature, and, since the nature of the human being is social, the virtuous life is also a life in service of the common good. To my mind, this is a naïve and insufficient view, because the social nature of the person is not enough to explain the common good as relational good (Donati 2019b). We need many intervening factors between what we call the nature of a human person and its social behavior. The factors become more and more complex as society becomes more and more 'relational,' i.e. oriented to free humans from natural limitations and constraints.

The idea that relations are a product of subjective consciousness (the first-person perspective) corresponds to what I call a naïve conception of the social as a human reality not mediated by what 'is among' human beings, according to which social relations *and their effects* are a sort of 'prolonging' or result of feelings (good or evil), of virtues (or vices), of an autonomous intimate life (or estrangement) of persons.

From the sociological standpoint, this derivation (induction) is problematic, if understood *sic et simpliciter*. Sobriety and the care of Creation, brotherhood, and other social virtues are now, in fact, in crisis precisely because it is no longer sufficient that the person wants them intentionally, whether it be a single person or a 'moral person', like a civil association or social movement.

There is something that lies ‘in the middle’ between the actions of single individuals to which we must give new focus. The social virtues, in fact, differently from the individual ones, refer to social relations. We call them both human, but is there any difference between them? Social virtues do not arise in an immediate and spontaneous way from within people, because, between the interior life of the person and social reality, they emerge (and are increasingly multiplying) from mediations (made up of relations) that make the immediacy and spontaneity of social outcomes completely uncertain and improbable.

In the new globalized environment, in order to create a certain social relation equipped with certain human qualities and causal properties (for example, in a couple, family, workplace, street, neighborhood, social networks, etc.), it is not only necessary to have a certain disposition and agency of individuals, but another condition becomes just as necessary: namely, that the persons ‘see’ the specific good (or evil) of that relation, and pursue that relation as good (or bad) in itself.

In the absence of this condition, the individual act, even the most virtuous and best intentioned, can create—albeit unintentionally—a relational evil instead of a relational good. Daily life is full of cases in which people who are in themselves good and close to each other create contentious and negative relationships. What went wrong? From a sociological viewpoint, what did not work was precisely the relation between them. While being good persons individually, they have neither seen nor cared for their relationship, which had decisive influences on themselves and the people around, whether or not they were aware of it.

The fact is that the social relation has its own reality, which is an ‘emergent’ not automatically derived from the qualities and dispositions of the individuals in relation. This is nothing other than what relational sociology says. From this standpoint the social relation appears as an enigma. In order to cope with it, I argue, we need a transcendental symbolic matrix.

We are dealing with exploring a new horizon, that of a culture of inter-human relations that can allow people to creatively respond to the inevitable enigmas of living together not by having recourse to ‘more and

more tech', but transcending the existing social context and human limits through a new relationality.

Conclusions: human beings transcend themselves in social relations, not technology

In this contribution, I have tried to confront the old humanism of the West, taking into account the deep changes of our times that are trying to transform the human into something else through technological progress. Basically, I see the possibilities of a neo-humanism in overcoming the old 'personalistic views' that have put what is human inside the individual as such rather than in its inner relational constitution.

The first-person perspective advocated by Lynne Rudder Baker (Baker 2015) claims that humans and non-humans have in common 'personhood' and differ by their bodies. She assumes that the unity between the material and the ideational elements of any entity has no constitutive relational identity. To me, this means assuming that *the nature of the relationship* between the human body and the human mind can be equated to the relationship between bronze material and a bronze statue. By this way we arrive at the result of considering smart robots as persons, precisely 'electronic persons', *endowed with individual rights and duties*.

This is where we arrive if we see electronic agents, as well as non-human animals, as new players in the political and legal arenas, as Gunther Teubner does (2006). If a robot were regarded as a legal entity, it could be held liable for its actions. Robot ethicists often consider a robot a moral actor, but their arguments are deeply questionable. Amitai and Oren Etzioni (2017) have argued with valid reasons that AI and robots have no moral agency, and that the term 'autonomous', commonly applied to these machines, is misleading, and leads to invalid conclusions about the ways these machines can be kept ethical. This is where we find the enigma of the social relationships that these 'persons' could activate.

Are the qualities and causal properties of the social relations acted by non-human 'persons' (as defined by Baker) equal to those proper to human

persons? Certainly not. The relation between organism and reflexivity in a human being has a level and kind of contingency and context-dependency that cannot have any functional equivalent in a machine, whatever its ability to act as a person ('I think I am I') in respect to the matter it is made of.

Recognizing the dilemma inherent to the body-mind relationship existing in a human being as different from that of a machine means seeing the proper qualities and causal power of human persons in respect to any other entity. No artificial intelligence can manage the enigmas of human relations as human beings do, although the latter can respond to the enigmas in the most absurd, violent and deviant ways. I don't think that machines might in principle be persons, because, although they can behave by reflecting in the first person, ontologically they lack the potential for expressing and managing proper relational feedbacks, relational generative mechanisms, and in particular 'relational steering' (Donati 2013). Possibly machines can be more predictable and controlled, but replacing human beings with machines would mean condemning human persons incapable of managing their enigmas to become servants of machines (Morozov, 2013). It would create a society where few people command and the masses were intended to be treated as scrap.

Trying to resolve the enigma in a human way means being able to configure the qualities and properties of the relations in such a way that those involved can share something essential while their differences bloom. Obviously, this does not always occur even in human beings, because it requires an adequate personal and relational reflexivity. What is sure is the fact that it cannot ever happen for a machine.

If one tackles the issue of hybrid liability of intelligent systems, one has to recognize the existence of what Werner Rammert has called 'distributed agency' (Rammert 2008). In distributed agency human action is distributed between many loci and instances that plan, control, and execute the activities. Agency is distributed between human beings. That is also the case in the hybrid liability constellation. Agency remains with natural persons. Robots work as mediators. Taken together, robot mediators have impact on the liability division between human beings.

Posthumans and transhumans are dreams populated by phantoms, and very peculiar phantoms that appear to be real, or better, seem to be real because they produce some realities. As Günther Anders (2016: section 24) claimed many years ago: “Phantoms are not only matrices of the experience of the world, but the world itself. The real as reproduction of its reproductions.” The kind of evolution that they feed is not human. To avoid all the deceits that they bring with them, we need a neo-humanism that leaves behind the old humanism.

Humanism is no longer synonymous with ‘personalism,’ since what is human and what is personhood are more and more differentiated (according to the relationist idea, put forward by Baker, that there can be persons who are not human). Still today traditional personalism distinguished humans from other living beings for their ability to exercise an inner reflexivity and make individual choices of communicated ultimate concerns, on the assumption that the goodness of these features would bring spontaneously the common good. Neo-humanism must acknowledge that this simple concept of the person is no longer sufficient to identify the human person. The latter should be redefined from a relational perspective. In the new scenario, a person is human, and becomes more human (that is transcends herself), if and to the extent that she generates social relationships that support the flourishing of relational goods from which she feeds herself in order to be more humane (Donati 2019c).

Some may come to the misgiving that my perspective affirms two contrary things: that relation is already there and at the same time must come into being. How does one state that “in the beginning there is the relation” if relation must be realized? Does this affirmation not have the Platonic flavor of supposing that there is an ideal subsistent reality that must only be actuated? This is not so. What I am arguing is the fact that persons come into existence and live in a context (which is always there) *from which* they are put into a structural relation within which they must act relationally. They are thus found necessarily in a morphostatic/morphogenetic process in which and through which people can reproduce preexisting relations or create new types of relations that confer on them previous identities or

change them. These emergent identities may or may not meet (in various ways) the potentiality of the persons, and from here the greater or lesser happiness of the persons, according to the sense of greater or lesser human self-realization.

The transcendence of what is human cannot but happen with reference to this relational constitution. We have to look at the ontology of social relationality, which has to maintain the autonomy of the human beings and at the same time bring out their relational good, in which they transcend themselves together with their differences. Trust, cooperation, caring for other people, and friendship are examples *par excellence*. Future highly intelligent and sophisticated machines will help human beings to transcend themselves, but transcending the human can only be a relational reflexive activity of human persons.

When we do not know what to do with others and with the situations of life, or what relation to have with the contingent world around us, then we feel confused, weak, fragile, sad, and in crisis. Every existential situation in which we find ourselves, each encounter with something or someone that puts us in trouble, is a relation that challenges us. Usually, we do not think of these situations in terms of relations, because we only see individuals and things. We must confront something (a situation) or someone (people around) and ask ourselves what to do. The fastest way to go is to find a technological device that can solve this problem. In reality, however, behind the challenge of situations, there is a challenge that we do not see, and it is the most important challenge. It is the challenge of the relation in play, which asks us to transcend ourselves.

We need to see the *Vital Relation* that, by linking the unity of body and mind to the context of existence, makes it possible for us to transcend ourselves. We must learn to deal with this *Vital Relation*. The enigma to be solved lies in this relation, not in technological devices. *Human life transcends itself when it sees the riddle of the relationship and finds the way out from it. The enigma of the relation contains the meaning of human life.*

My answer to the question of why, when and how the human can be transcended is therefore the following. The crisis of human relations is

what fuels the search for a posthumanism that can remedy this crisis. The concomitant crisis of a collective faith and of interpersonal relations that we witness in contemporary advanced societies finds an explanation in the fact that in both of these relations the substance of relationality is diminished, that is, the qualities and causal properties of the reciprocity between the two terms of the relation are lacking. The loss of the inherent sense of the relation is the ‘intervening factor’ which explains both of the crises. If the sense of the human relation diminishes, then so does the transcendental relation, and vice versa. The same holds true for the relation between the human and infrahuman realms. For a deeper understanding of this point, it is necessary to recall that the source of the *meaning of the relation* lies in its transcendental matrix, although the social relation is generated (or is not generated) due to autonomous causes.

In summary, to grasp the why, where and how the human can transcend itself, it is necessary to ‘see’ the peculiarities of interpersonal relations and their enigmas. One could observe that even the relations with robots are enigmatic, but these enigmas are games, not the stakes of serious life (the Durkheimian “*vie serieuse*”). Thinking that the human person is distinguished from animals, plants and machines because it is able to think in first-person (i.e. to think of oneself as oneself) is not enough. Nobody can exclude the possibility that, in the future, very smart robots or cyborgs could become able to behave in such a way. Calling them ‘persons’ as distinct from ‘human persons’ means to go back to the Greek concept of person as a mask, for example in the modern version of puppets proposed by Goffman, or, in a more sophisticated way, it means to embrace the Luhmannian (neo-functional) view of the human person as a mere ‘reference point for communication,’ which removes the distinction between human and non-human.

To my mind, what distinguishes the human person is her ‘*structural relational constitution*,’ in the inner as well as exterior life. This dynamic constitution is the relational substance that enables the human person to transcend herself in the relationship with others and with the world, acting reflexively, not only in herself (personal reflexivity), but on social relations as such (relational reflexivity), which she assesses as good or evil, taking into

account the behaviour of significant others and reference to the situational context. The first-person perspective that has supported the traditional humanism for centuries should be included and upgraded within a new relational paradigm, otherwise the humanism that it seems to defend will be just the latest attempt to keep alive the old Western individualism vis-à-vis the arrival of ‘enhanced humans,’ androids, robots and cyborgs.

The technologies that lead humanity towards the post/trans-human must be analyzed and evaluated based on the criteria of which human relations they assume, and of those they produce. We have to see if they support arrangements that increase the ethical and empathetic sense of social relations or, vice versa, nurture relationships that are empty of human sense.

References

- Anders, Günther. 2016. *Prometheanism: Technology, Digital Culture and Human Obsolescence* [1956] New York: Rowman & Littlefield.
- Archer, Margaret. 2019a. “Bodies, persons and human enhancement: Why these distinctions matter.” In I. Al-Amoudi and J. Morgan (eds.). *Realist Responses to Post-Human Society: Ex Machina*. Abingdon: Routledge, pp. 10–32.
- Archer, Margaret. 2019b. “Can Human Beings and AI Robots be Friends?.” Paper delivered at the PAS-PASS conference “Robotics, AI, and Humanity: Science, Ethics, and Policy”, Vatican City: May 16–17.
- Baker, Lynne Rudder. 2002. “The Ontological Status of Persons.” *Philosophy and Phenomenological Research* 65: 370–388.
- Baker, Lynne Rudder. 2007. “Naturalism and the First-Person Perspective.’ in *How Successful is Naturalism?*.” Frankfurt: Ontos-Verlag, pp. 203–226.
- Baker, Lynne Rudder. 2013. “Technology and the Future of Persons.” *The Monist* 96: 37–53.
- Baker, Lynne Rudder. 2015. “Human Persons as Social Entities.” *Journal of Social Ontology* 1(1): 77–87.
- Baker, Lynne Rudder. 2016. “Making sense of ourselves: self-narratives and personal identity.” *Phenomenology and the Cognitive Sciences* 15: 7–15.
- Bilimoria, Purushottama. 2012. “Why Is There Nothing Rather Than Something? An Essay in the Comparative Metaphysic of Nonbeing.” *SOPHIA* 51:509–530.

- Bhaskar, Roy. 1989. *The Possibility of Naturalism*. Hemel Hempstead: Harvester Wheatsheaf.
- Bhaskar, Roy. 2012. *Reflections on MetaReality: Transcendence, Emancipation and Everyday Life*. London: Routledge.
- Breslau, Daniel. 2000. "Sociology after humanism: A lesson from contemporary science studies." *Sociological Theory* 18 (2): 289–307.
- Cath, Corinne et al. 2017. "Artificial Intelligence and the 'Good Society': the US, EU, and UK approach." *Sci. Eng. Ethics* published on line 28 March 2017. DOI 10.1007/s11948-017-9901-7
- Debaise, Didier. 2012. "What is Relational Thinking?." *Inflexions* 5 : 1–11.
- Donati, Pierpaolo. 2011. *Relational Sociology. A New Paradigm for the Social Sciences*. Abingdon: Routledge.
- Donati, Pierpaolo. 2013. "Morphogenesis and Social Networks: Relational Steering not Mechanical Feedback." In M.S. Archer (ed.). *Social Morphogenesis*. Dordrecht: Springer, pp. 205–231.
- Donati, Pierpaolo. 2019a. *The Enigma of Relation and the Theological Matrix of Society*, in P. Donati, A. Malo and G. Maspero (eds.). *Social Science, Philosophy and Theology in Dialogue: A Relational Perspective*. Abingdon: Routledge, pp. 3–37.
- Donati, Pierpaolo. 2019b. "Discovering the Relational Goods: Their Nature, Genesis and Effects." *International Review of Sociology/Revue internationale de Sociologie* DOI 10.1080/03906701.2019.1619952
- Donati, Pierpaolo. 2020. *The Digital Matrix and the Hybridization of Society*, in I. Al-Amoudi and E. Lazega eds. *Post-human institutions and organisations: confronting the Matrix*. Abingdon: Routledge.
- Donati, Pierpaolo. 2021. "Being Human in the Digital Matrix Land." In M. Carrigan, D. Porpora, and C. Wight (eds.). *The future of the human and social relations*. Abingdon: Routledge (in print).
- Donati, Pierpaolo and Archer, Margaret. 2015. *The Relational Subject*. Cambridge: Cambridge University Press.
- Etzioni, Amitai and Etzioni, Oren. 2017. "Incorporating Ethics into Artificial Intelligence." *The Journal of Ethics* 21(1): 1–16.
- Floridi, Luciano. ed. 2015. *The Onlife Manifesto. Being Human in a Hyperconnected Era* New York: Springer.
- Gehlen, Arnold. 1974. *Der Mensch. Seine Natur und seine Stellung in der Welt* [1940]. Frankfurt am Main: Athenaion Verlag.
- Jones, Raya A. 2013. "Relationalism through Social Robotics." *Journal for the Theory of Social Behaviour*, 43 (4): 405–424.

- Jullien, François. 2015. *De l'être au vivre. Lexique euro-chinois de la pensée*. Paris: Gallimard.
- Luhmann, Niklas. 1995. *Social Systems*. Stanford: Stanford University Press.
- McCormack, Mark, Wignall, Liam and Morris, Max. 2016. "Gay guys using gay language: friendship, shared values and the intent-context-effect matrix." *The British Journal of Sociology* 67(4): 747–767.
- Morozov, Evgenij. 2013. *To Save Everything, Click Here. The Folly of Technological Solutionism*. Philadelphia: PublicAffairs.
- Parsons, Talcott. 1937. *The Structure of Social Action*, New York: Free Press.
- Rammert, Werner. 2008. "Where the action is: Distributed agency between humans, machines, and Programs." Bielefeld: The Technical University Technology Studies Working Papers TUTS-WP-4.
- Silver, Daniel and O'Neill, Kristie. 2014. "The significance of religious imagery in *The Philosophy of Money: Money and the transcendent character of life*." *European Journal of Social Theory* 17(4): 389–406.
- Teubner, Gunther. 2006. "Rights of Non-humans? Electronic Agents and Animals as New Actors in Politics and Law." *Journal of Law and Society* 33(4): 497–521.
- Weber, Max. 1993. *The Sociology of Religion* [1920]. Boston: Beacon Press.