

New Moral Subjects: Walking the Line Between Autonomy and Automaton

Ralf Lüfter

Free University of Bozen-Bolzano, Bozen, Italy

The question of “new moral subjects” arises today in connection with the use of advanced technologies in the context of Industry 4.0, i.e., the industrial enterprise investing in smart factories, co-robotic, additive manufacturing, augmented reality, cyber security, big data, etc. Industry 4.0 underpins what is being heralded as Industry 5.0. Technologies in this field excel in that they are not used for the mere execution of processes designed, programmed, and controlled by man, but are applied to realise “autonomously” goals in the industrial context. Consequently, their degree of automation is no longer defined by increasing independence from human beings, but by the ability to spontaneously initiate a process and implement it to useful effect. The comparison between man and machine, and the justification for the introduction of “new moral subjects” thus crucially depend on the location of “moral being” in the reference to their “autonomy” as a property characterising the “subject man”, but not the exclusive prerogative of the latter.

Keywords: new moral subjects, ethics of automation, responsibility, autonomy, Industry 4.0

Introductory Remarks

The aim of this short paper is not—and cannot be—to present a complete line of argumentation that allows conclusions to be drawn about the question of “new moral subjects”. Rather, the following is a provisional and fragmentary attempt to outline some questions that arise *regarding the way in which the “new moral subjects” are thematised in the context of “advanced technologies” whose economic relevance is evident not least in the so-called “Industry 4.0” (and the subsequent “Industry 5.0”)*. These questions are intended to sketch a field of research that lies idle to some extent whereas it might become a focus of future ethics. One of these questions is the question of “autonomy”, or more precisely: the question of “man’s autonomy” in the “age of technics”.

Here we are already touching on a problem which at first glance is a purely terminological one—but which ultimately goes beyond terminological distinctions, and which returns as a problem towards the end of this presentation. This is to say that for the present discourse, we must distinguish between “technology” and “technics”.

The field of “technology” is composed of *technological objects* that are distinguished from “ordinary” objects by virtue of their peculiar characteristics, and those problems that arise in their application. There is, of course, a relevant concern about technology, about positive and negative consequences of the use of technology,

The article is the elaborated and extended version of a presentation given at the 30th Annual International Vincentian Business Ethics Conference, at Dublin City University in October 2023.

Ralf Lüfter, Ph.D., assistant professor, Faculty of Economics and Management, Free University of Bozen-Bolzano, Bozen, Italy.

about the way we should handle technology, about the human control of technology, about man in turn being influenced or even controlled by technology, and so on.

On the other hand, the notion of “technics”, understood from the Greek word τέχνη, refers to a way of “letting things be”, i.e., of bringing them about, of bringing them to light, and thus does not present itself as a concept of causing or effecting something; on the contrary, this notion presents itself as a concept of understanding, of knowing, of “knowing how to do” in the sense of knowing how to discern what comes about, what comes to light. And this refers to another field of concern (different from the one related to technology and its use). We will come back to this at the end of the presentation (Heidegger, 1989; 2002).¹

Thematic Context

Today, the question of the “new moral subjects” arises not least in relation to the use of so-called “advanced technologies” in the context of Industry 4.0, i.e., an industry that invests in *smart factories*, *co-roboting*, *additive manufacturing*, *augmented reality*, *cyber security*, *big data*; this way of thinking industry—the “Industry 4.0”—is in turn the basis of what is already being heralded as “Industry 5.0”, i.e., the vision of a more “human-friendly” industry, which integrates “productivism” (i.e., the increase in productivity through the involvement of advanced technologies) and “sustainable development” (i.e., the transition to stable economic growth that is socially acceptable and respects the ecological balance) (Miller & Wendt, 2021; Hooker & Kim, 2019; Schwab, 2017; Elliot & Spencer, 2017).

The technologies (advanced technologies) applied in this context *excel to the extent* that they are not employed in the mere execution of processes conceived, programmed, and controlled by man but, rather “act”, so to speak, “autonomously” in the sense of implementing those goals that the industry aims to realise in the first place and that therefore are in its interest.

Consequently, the degree of automation of such technologies is no longer defined by their increasing independence from man, but by their ability to spontaneously initiate a process and execute it to the point of producing useful effects. In fact, according to Daniela Tafani (2023), it should be noted that the “autonomy” of technologies employed in the context of “Industry 4.0” is not primarily constituted by their increasing independence from human intervention and control, but “by the fact that these technologies appear to be able to perform specific tasks, which had been, until now, the prerogative of human beings alone”.

In other words, in this context, the understanding of “autonomy” is constituted in relation to the degree of automation that is realised through these specific technologies, that is, as we shall see in a moment, by the extent to which they seem to be able to perform typically human actions while the degree of automation *is in* turn based on the concomitance of three factors: the availability of data (we are dealing with an exponential increase in data that is processed in such a way that it can be used by the “advanced technologies” within the framework of computing operations), the optimization of computation, and eventually (as a condition for the possibility of improving computation) the increasing functionality of algorithms.

Following the argumentation developed so far, the linking between “autonomy” and “automation” is established in this context in such a way that the notion of “autonomy” seems to be defined by the “degree of automation” of the technological objects in question.

¹ The understanding of “technics” (and its distinction from “technology”), proposed here, is suggested by Martin Heidegger.

In short, the notion of “autonomy” is traced back (we are not sure by which routes) to the concept of “automaton”.

For we call that which “moves by itself” “automaton”, and we distinguish between, on the one hand, those “automata” that merely imitate a “moving by itself” (and are therefore not at all autonomous, since they are mere machines that mimic movements mechanically and present an anthropomorphic image of movement)², and, on the other hand, those “automata” whose “moving by itself” is generated as striving (for something), as aiming (at something) and who through “moving by itself” can put into action what they aspire to and aim at. The “autonomy” of the latter would thus become manifest in their ability to decide the end for which they themselves—“moving by themselves”—are means. “Autonomy” would thus be a property (their property), i.e., the property of “proposing” an end for the implementation of which the “proponent” of the end is at the same time a means of its (the end’s) realisation.

The Automaton as Autonomous Agent

At this point we must ask ourselves: What kind of “autonomy” are we talking about? Is there an analogy between *the autonomy of the automaton* and *the autonomy of the human being*? From what, if anything, can we derive such an analogy? Or have we reached the point where we admit that man himself is conceivable as an autonomous automaton in the sense just described?

This goes hand in hand with the following circumstance: The “advanced technologies” applied by the “Industry 4.0” challenge the ordinary conception of technology as such—namely, technology as a tool, as an instrument that man uses to achieve ends that he himself sets and realises by the virtue of his free will. According to this conception, technology belongs to the realm of human action, and thus falls under human responsibility, both in terms of its use and the effects caused by this use (Jonas, 2010).

The “advanced technologies” underlying “Industry 4.0”—as assumed here—differ from the ordinary conception of technology in that they are no longer mere objects in the sense of instruments of human use and would therefore no longer fit into *the anthropological-instrumental conception of technology*; rather, these technologies are seen as subjects of action, that can make “autonomous” decisions and can consequently be held responsible (M. Anderson & S. L. Anderson, 2009).

Recent research has made important contributions on this issue, focusing, however, mainly on understanding the differences and similarities between the human being and machines, both of which are in their own way understood—despite the gap that still seems to separate the human being from machines—as “autonomous agents” (Bendel, 2019; Misselhorn, 2018).

“Old Moral Subjects” and “New Moral Subjects”

But is there now any reason on the ground of which we could assume that there are, on the one hand, “old moral subjects” and, on the other hand, “new moral subjects”? Or does the notion of “moral subject” itself become increasingly instable and challenging in this context and therefore needs to be questioned further?

² In the 18th century, the word “automaton” was used to designate machines that were able to imitate the movements and often even the outward appearance of men and animals (e.g., Pierre Jaquet-Droz’ machines). Ottorino Pianigiani’s etymological vocabulary reads under the heading “automaton”: “machine that moves by hidden devices, so that it seems to move by itself, as if it had life” (Pianigiani, 2023).

Traditionally the notion of “moral subject” is strongly linked to that of “autonomy”: The subject is moral insofar as it is able to respond spontaneously to the need to “give itself its own norm (its own measure)” and to conform to it through its actions.

Following the traditional conjecture, “morality” consists (not least) in this distinct property of the subject, which justifies the introduction of “new moral subjects” if any entities (in this context: technological entities) prove to be endowed with analogous properties (or to be endowed with the capacities to develop analogous properties).

Now, if the conceptualization of the “new moral subjects” remains on the level of the meant analogy—on the level of differences and similarities between human beings and machines—without questioning the consistency of the analogy itself for its soundness, it remains limited, from the outset, to the assumption that “new moral subjects” merely expand (or rather do not expand) the traditional circle of moral subjects, depending on whether we guess that they add to (or do not add to) the latter, and that they break (or do not break) the exclusivity of the previous (old) moral subjects.

On the level of this analogy the question does not arise as to whether the notion of “moral subject” itself may be sufficient in order to understand “autonomy” as one of the fundamental traits of a future ethics, if it is true that, as Hans Jonas argues, the new territory of the collective practice, which we have entered with high technology (to which also the “advanced technologies” belong) is still a no-mans-land for ethical theory (Jonas, 2010).

The comparison between man and machines, and the justification for the introduction of “new moral subjects” therefor depend on the location of “being moral” in reference to “autonomy” as a property that characterises the “human subject”, without, however, being the exclusive prerogative of the latter.

On the other hand, the assumption that the notion of “autonomy” can be separated from the “human being” contradicts what is proposed by the philosophical tradition, especially Kant, who, for example, sees in the public use of reason a way out of a state of minority (of immaturity), insofar as the essence of the human being lies precisely in his ability to respond autonomously to the obligation to sustain and preserve his own being and thus to exist as a human being in the first place. This ability, which is founded in man’s autonomy and located in his subjectivity, distinguishes him from all other entities (Kant, 2009).

In this context, questions like this arise: In what sense can autonomy be conceived as a property of the moral subject? In what relation do autonomy and morality stand? How is it that autonomy seems to be a condition of possibility for the morality of subjects of action?

Man’s Original Responsibility

The criticality of the conceptualisation of “new moral subjects” increases even more, when the analogy between man and machine is abandoned and the notion of “autonomy” is not conceived automatically as a property of man (which in turn is viewed exclusively in the light of subjectivity) that can be applied by analogy to technological objects but is itself taken as a source for conceiving the human being as such. In this perspective, “being moral” consists in an original responsibility towards “autonomy” that calls (obliges) man to his own humanity.

At this point we can return to the distinction, which was introduced at the beginning: With reference to “the human being” in “the age of technics” we addressed the difference between “technology” and “technics”—between “technological objects” and “a way of letting things be” (whereas the latter refers to Greek word τέχνη).

Now, precisely in the light of this difference, “technics” appears as a configuration of man’s original responsibility: In fact, in as much as “technics” is thought as “a way of letting things be” by “knowing how to do”, it is itself considered as a way of taking care of “the sense” to be brought to light. Such taking care is considered as an original responsibility insofar as it responds to the call of sense itself, i.e., to the need (obligation) to take care of it, on pain of the devastation of those sense-relations in which man exists, and consequently the devastation of the world as the “dimension of the human dwelling”—literally speaking, of the $\tilde{\eta}\theta\omicron\varsigma$.

Within the occidental tradition of thinking, the question of ethics is born out of that sense-dimension which requires to be sustained in and through the being of man, in view of the foundation of his abode, to wit his $\tilde{\eta}\theta\omicron\varsigma$ (Liddell & Scott, 1996, p. 766). Ethics, in turn, originates as knowledge of “the dwelling of the human being, [as knowledge of] his abode in the midst of beings in the whole (i.e. the sense-dimension)” (Heidegger, 1994, p. 214).

So if we assume that what we are envisaging here is reliable, we can question the way in which we respond to the call of sense today (how the human being inhabits the sense-dimension today)—a way that is manifest, not least, in the “advanced technologies” and their use, as well as in the above-mentioned conjunction that places the human being in an analogy with the automaton. It is precisely this questioning that leads—with as yet unforeseeable consequences—to a rethinking of the concept of the “moral subject” in all its complexity, just as the “moral subject” itself runs the risk of being thought of as a kind of automaton, i.e., as the automaton whose actions are characterised by the property of morality.

References

- Anderson, M., & Anderson, S. L. (2009). *Machine ethics*. Cambridge: Cambridge University Press.
- Bendel, H. (Ed.). (2019). *Handbuch Maschinenethik*. Berlin: Springer.
- Elliot, D., & Spencer, E. H. (2017). *Ethics for a digital era*. Hoboken: Wiley.
- Heidegger, M. (1989). *Überlieferte Sprache und technische Sprache*. St. Gallen: Erker.
- Heidegger, M. (1994). *Heraklit*. HGA 55. Frankfurt am Main: Vittorio Klostermann Verlag. (My translation)
- Heidegger, M. (2002). *Die Technik und die Kehre*. Stuttgart: Klett-Cotta.
- Hooker, J., & Kim, T. W. (2019). Ethical implications of the fourth industrial revolution for business and society. *Business Ethics*, 3(7), 35-63.
- Jonas, H. (2010). *The imperative of responsibility. In search of an ethics for the technological age*. Chicago & London: The University of Chicago Press.
- Kant, I. (2009). *An answer to the question: What is enlightenment?* New York: Penguin.
- Liddell, H. G., & Scott, R. (1996). *Greek-English lexicon*. Oxford: Calderon Press.
- Miller, K., & Wendt, K. (2021). *The fourth industrial revolution and its impact on ethics. Solving the challenges of the agenda 2030*. Berlin: Springer Nature.
- Misselhorn, C. (2018). *Grundfragen der Maschinenethik*. Ditzingen: Reclam.
- Pianigiani, O. (10.12.2023). *Vocabolario etimologico della lingua italiana*. Retrieved from <http://etimo.it>
- Schwab, K. (2017). *The fourth industrial revolution*. London: Penguin.
- Tafari, D. (03.12.2023). Sistemi fuori controllo o prodotti fuorilegge? La cosiddetta “intelligenza artificiale” e il risveglio del diritto. *Bollettino telematico di filosofia politica*. Retrieved from <https://btf-sp-unipi.it>