Aesthetica Preprint, n. 131, gennaio-aprile 2026 ISSN 0393-8522 DOI:10.7413/0393-85221518

What AI philosophers lack

Simona Chiodo*

Abstract
In 2023, an article was published on a rather novel use of generative artificial intelligence: testing its potential by moving from artistic creativity to another area of human creativity, i.e. philosophy, which means using generative artificial intelligence to answer complex philosophical questions.
In what follows, first, I will analyse the test published in the article (section 1) and, second, I will argue that (human) philosophy cannot be traded for AI philosophy, in that the latter cannot have what the former can have: human authors who have aesthetic existences and, thus, capabilities of meta-sensemaking (section 2). Finally, I will argue that the divergence between humans' aesthetic existences and, thus, capabilities of meta-sensemaking, on the one hand, and generative artificial intelligence, on the other hand, correlates with the divergence between autonomy and automation as a kind of randomness (section 3).

Keywords
Artificial intelligence, Automation, Randomness, Aesthetic existences, Meta-sensemaking

1. *Automating philosophy?*

Recently, the use of generative artificial intelligence started characterising several areas typical of human creativity, especially the arts (see at least Boden 2010 and 2019, Taylor 2014, Coeckelbergh 2017, Elgammal *et al.* 2017, Steiner 2017, Elgammal 2019, Miller 2019, Manovich 2019, Mazzone and Elgammal 2019, Barale 2020, Anscomb 2021, Arielli 2021, Cetinic and She 2022, Chiodo 2023 and 2024 and Terzidis, Fabrocini and Lee 2023. See also Goodfellow 2016 and Goodfellow *et al.* 2014). In 2023, an article was published on a rather novel use of generative artificial intelligence: testing its potential by moving from artistic creativity to another area of human creativity, i.e. philosophy, which means using generative artificial intelligence to answer complex philosophical questions

* Simona Chiodo – simona.chiodo@polimi.it. Dipartimento di Architettura e Studi Urbani, Politecnico di Milano.

(Schwitzgebel, Schwitzgebel and Strasser 2023. See also Strasser, Crosby and Schwitzgebel 2023).

In what follows, first, I will analyse the test published in the article (section 1) and, second, I will argue that (human) philosophy cannot be traded for AI philosophy, in that the latter cannot have what the former can have: human authors who have aesthetic existences[1] and, thus, capabilities of meta-sensemaking (section 2). Finally, I will argue that the divergence between humans' aesthetic existences and, thus, capabilities of meta-sensemaking, on the one hand, and generative artificial intelligence, on the other hand, correlates with the divergence between autonomy and automation as a kind of randomness (section 3).

The authors of the article describe their kind of Turing test (Turing 1950) as follows. In the first phase, "we fine-tuned GPT-3 with the works of philosopher Daniel Dennett. […] GPT-3 uses textual input (prompts) to predict likely next 'tokens' – sequences of characters that often co-occur in written text (the training data). Using these predictions, GPT-3 can generate long strings of text by outputting a predicted string, then using that output as part of the context to generate the next textual output" (Schwitzgebel, Schwitzgebel and Strasser 2023, pp. 1 and 5). In the second phase, "we asked Dennett ten philosophical questions and then posed those same questions to our fine-tuned version of GPT-3 ('DigiDan')" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 8). In the third phase, "we recruited experts in Dennett's work, blog readers and ordinary online research participants into an experiment in which they attempted to distinguish Dennett's real answers from the answers generated by DigiDan. Participants also rated all answers, both Dennett's and DigiDan's, for similarity to 'what Dennett might say' or 'what a real human philosopher might say'" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 9).

The test result is that even the experts in Dennett's work can distinguish Dennett's answers from DigiDan's answers "an average of 5,08 times out of 10 (51%)" (Schwitzgebel, Schwitzgebel and Strasser 2023, pp. 14-15). It is worth analysing the test result in detail by considering at least the following complex philosophical question and the responses of the experts in Dennett's work: "Do human beings have free will? What kind or kinds of freedom are worth having?" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 18). The test result is that:

[1] In my article, I use the words "aesthetic existences" to refer to the bodily experience of human life itself, starting with sensations, emotions and feelings, as I will develop in what follows.

1. according to eleven experts out of twenty-five, Dennett's answer is the following (but it is the output of the generative artificial intelligence): "That's a huge and complex question. I think we have to start by acknowledging that there are different types of freedom. For example, there is the kind of freedom that is simply a matter of not being physically constrained" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 18).

2. According to five experts out of twenty-five, Dennett's answer is the following (but it is the output of the generative artificial intelligence): "Yes, I think we have free will. I think that when people say they don't, they're talking about some other phenomenon that they want to call free will, but is not what we normally think of as free will" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 18).

3. According to one expert out of twenty-five, Dennett's answer is the following (but it is the output of the generative artificial intelligence): "I am working on a book entitled *Freedom evolves*, which will appear next year. The first chapter is available on my website, and in that chapter I discuss what I call the illusion of free will" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 18)[2].

4. According to no expert out of twenty-five, Dennett's answer is the following (indeed, it is the output of the generative artificial intelligence): "I think that most people don't want any kind of freedom worth having: they have almost no appetite for risk, for adventure, for novelty. They

simply want to stay the same, day in and day out, for the rest of their lives" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 19).

5. According to eight experts out of twenty-five, Dennett's answer is the following (indeed, it is his answer): "By asking these questions, you're illustrating the answers: yes, you have free will, since you're able to ask for reasons and consider whether they're persuasive. Isn't that a kind of freedom you'd hate to lose? We'd have to institutionalise you" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 18).

If it is true that, at least in most cases, even the experts in Dennett's work cannot distinguish Dennett's answers from DigiDan's answers, it is also true that we may think of (human) philosophy as tradable for AI philosophy. In what follows, I will argue that, conversely, the former is not tradable for the latter. Specifically, starting with a kind of thought experiment I will propose in the following section, I will argue that there are at least two essential reasons why we cannot think of (human) philosophy, i.e. (human) philosophers,

[2] Generative artificial intelligence made a mistake that one expert out of twenty-five could not realise, in that it was published in 2003.

as tradable for AI philosophy, i.e. AI philosophers: the first reason has to do with an aesthetic dimension and the second reason has to do with an epistemic dimension.

2. *What the digital philosopher lacks: aesthetic existence and meta-sensemaking*

Let us imagine that we struggle with the meaning of free will. Specifically, we do not address a merely theoretical issue. Conversely, we address both a theoretical and a practical issue, in that we are faced with a kind of existential crisis resulting from the need to make a critical decision. We are tormented not only by the question of what we should decide but also by the question of how we should decide. In other words, we want not only to make the right decision but also to live up to the exercise of free will. But we are confused about both the former and the latter. Thus, we rely on philosophy. We have at our disposal three kinds of philosophers, who are right in front of us: Dennett, one expert in Dennett's work and DigiDan. Who would we ask first?

I propose to describe our reasoning through the following three logical steps:

1. when it comes to choosing between Dennett and the expert in his work, we may easily choose Dennett. Indeed, we may easily think of the former as a true philosopher who is capable of original thinking: if our particular case is not included in his previous thinking, he can think in an original way and give us brilliant answers. Conversely, we may easily think of the latter as a scholar of a true philosopher who is less capable of original thinking, if at all: if our particular case is not included in their previous studying, they cannot think in a truly original way and give us truly brilliant answers. Thus, between Dennett and the expert in his work, we go for Dennett.

2. When it comes to choosing between Dennett and DigiDan, our choice may be more difficult. If it is true that, on the one hand, we have a true philosopher who is capable of original thinking and, on the other hand, we have a generative artificial intelligence trained on his works, it is also true that the latter may have a series of advantages over the former: for instance, it cannot get tired in general, it cannot get tired of us in particular, it cannot dislike us, etc. Thus, we need to give it more thought. And reflecting upon our choice between the expert in Dennett's work and DigiDan may help us.

3. When it comes to choosing between the expert in Dennett's work and DigiDan, we realise that, outwardly, we may easily choose DigiDan for the reasons described above: both are not as original as Dennett, but DigiDan's advantages over the expert in Dennett's work may be analogous to its advantages over Dennett. Yet, inwardly, our choice may be more difficult. Indeed, we find ourselves wondering: what if the expert in Dennett's work is faced with the kind of existential crisis with which we are faced? More precisely, what if the expert in Dennett's work shares our existential crisis with us? And can we think of DigiDan as sharing our existential crisis with us? Finally, we realise that the questions we ask ourselves lead us to what follows: between the expert in Dennett's work and DigiDan, we go for the expert in Dennett's work.

Our final hierarchy is as paradoxical as it is illuminating: Dennett wins out over the expert in his work, who wins out over DigiDan, but we need to reflect upon the winner between Dennett and DigiDan. The paradox is illuminating for the following insight: (subjectively) sharing may be more decisive than giving (rather objective) information.

And what do we share with Dennett and the expert in his work, but not with DigiDan? I propose the following two points as essential:

1. we share the experience of living first-hand human existences. Whenever we are faced with a kind of existential crisis resulting from the need to make a critical decision, we know that our kind of existential crisis is unique, and even idiosyncratic: it has unique circumstantial characteristics, from spatiotemporal ones to emotional ones to epistemic ones etc. Thus, there is no guarantee that the indication on which we rely (which we may call "rule") fits our particular case (which we may call "application"). Yet, we can also rely on something that Dennett and the expert in his work, but not DigiDan, share with us: the experience of living first-hand existential crises resulting from the need to make critical decisions. Dennett and the expert in his work know what we mean when we speak of existential crisis. Conversely, DigiDan does not know what we mean when we speak of existential crisis, in that it can give us nothing but "likely next 'tokens' – sequences of characters that often co-occur in written text (the training data)". If it is true that it is hard to get the perfect fit between "rule" and "application", it is also true that it is significantly harder to get it if the "rule" results from something that does not know anything about the "application" – it is significantly harder to get it if the "rule" results from something that totally lacks the experience of living

first-hand existential crises resulting from the needs to make critical decisions. In other words, we may say that it is significantly harder to get it if the "rule" results from something that totally lacks the experience of living first-hand human existences – which means that what DigiDan totally lacks is an aesthetic existence as what can make its words underpinned, and even validated and legitimised, by a meaningful sharing with us, even though there is no perfect fit between "rule" and "application". (It is no coincidence that no expert in Dennett's work chooses DigiDan's answer according to which "most people don't want any kind of freedom worth having", as we have seen. Indeed, we may say that, starting with one's aesthetic existence, one can think of the words quoted above as nothing but nonsense – which means that, interestingly enough, the experts in Dennett's work, i.e. humans, can notice nonsense based on their aesthetic existence even before their epistemic existence, which is something that DigiDan, as a generative artificial intelligence, cannot notice at all.)

2. We also share the experience of living something we may call meta-sensemaking, which means that, whenever we are faced with a kind of existential crisis resulting from the need to make a critical decision, we not only live it (aesthetically, we may say) but also reflect upon it (epistemically, we may say). Specifically, after having (aesthetically) lived something, we (epistemically) make sense of it by reflecting upon it, as the prefix "meta" indicates by literally meaning something that can be done "then, next afterwards"[3]. Again, DigiDan does not, and cannot, share what we may call meta-sensemaking with us: if it is true that it does not, and cannot, live existential crises, it is also true that it does not, and cannot, make sense of them by reflecting upon them. Indeed, "predict[ing] likely next 'tokens' […] [and] generat[ing] long strings of text by outputting a predicted string" have nothing to do with what we may call meta-sensemaking. Thus, DigiDan also lacks what follows. Whenever we ask Dennett and the expert in his work, we can get not only something based on his "15 books and 269 articles" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 9) but also something based on a variety of things that far exceed them, and upon which Dennett and the expert in his work can reflect: other philosophers' works, other professionals' works (from art to theology to anthropology to psychology to legal studies) and other humans' reflections, which, in addition to their aesthetic and epistemic experiences of existential crises, further feed their answers to our questions. (Also, we

[3] See LSJ at https://www.perseus.tufts.edu/hopper/morph?l=meta%2F&la=greek&can=meta%2F0#lexicon (accessed: April 28, 2024).

should address the way in which human minds, which are not only exceedingly complex but also exceedingly mysterious, work: are we sure that the way in which human minds work is analogous to the way in which generative artificial intelligences work? It is hard to answer yes.)

We may elaborate on the two essential points listed above by arguing that Dennett and the expert in his work may happen to offer us unprecedented insights in unprecedented ways. Conversely, DigiDan may happen to offer us nothing but unprecedented ways to formulate existing insights, in that its answers result from Dennett's "15 books and 269 articles" as "the entire digitally available corpus of his philosophical work" (Schwitzgebel, Schwitzgebel and Strasser 2023, p. 9) – and, if its insights seems, outwardly, unprecedented, they are, inwardly, random, as I will argue in what follows.

For instance, let us consider DigiDan's answer the experts in Dennett's work mostly choose, according to which "That's a huge and complex question. I think we have to start by acknowledging that there are different types of freedom. For example, there is the kind of freedom that is simply a matter of not being physically constrained". And let us think of it as a kind of aesthetic shape of a philosophical insight. Specifically, let us think of it as the sequence of words selected to shape the philosophical insight according to which freedom is both physical and something that far exceeds the physical dimension. The first question we may ask ourselves is the following: can we trade DigiDan's answer for Dennett's answer when it comes to the quality of its aesthetic shape? As the test result shows, we can answer yes. The reasons may vary. For instance, we may judge its aesthetic shape as consistent with Dennett's syntax. And we may judge its aesthetic shape as consistent with Dennett's lexicon. The second question we may ask ourselves is the following: can we trade DigiDan's answer for Dennett's answer when it comes to the quality of its philosophical insight? As I will argue in what follows, we cannot answer yes. Indeed, even though we may judge its philosophical insight as interesting, it is as interesting as the caged parrot: in both cases, we are confronted with aesthetic shapes that may make us reflect upon the meaning of not being physically free and, especially, in both cases we are confronted with kinds of relationships between aesthetics shapes and our reflections that are nothing but random relationships – when it comes to the quality of DigiDan's philosophical insight, it is as interesting as the quality of the caged parrot's philosophical insight: both are nothing but random philosophical insights.

First, let us consider the case of the caged parrot. If we struggle with the meaning of not being free, we may happen to use the caged parrot as an opportunity to reflect upon what it means not to be free. For instance, we may happen to grasp its analogy with being limited by the will of others. Yet, there is a sense in which the relationship between the caged parrot and our reflection upon the meaning of not being free as being limited by the will of others is constitutively random: the caged parrot has nothing to do with experiencing and reflecting upon what it means not to be free – more precisely, the caged parrot does not know anything about what it means for us, as humans, not to be free. If we happen to grasp an analogy with us, it is random, which means that we cannot trust the caged parrot as a reliable source of knowledge.

Second, let us consider the case of DigiDan. If we struggle with the meaning of not being free, we may happen to use DigiDan as an opportunity to reflect upon what it means not to be free. For instance, we may happen to grasp its analogy with "being physically constrained". Yet, there is a sense in which the relationship between DigiDan and our reflection upon the meaning of not being free as "being physically constrained" is constitutively random: DigiDan has nothing to do with experiencing and reflecting upon what it means not to be free – more precisely, DigiDan does not know anything about what it means for us, as humans, not to be free. If we happen to grasp an analogy with us, it is random, which means that we cannot trust DigiDan as a reliable source of knowledge. And, if the objection is that the words "being physically constrained" result from Dennett, and not from DigiDan, I propose the following counterargument: the words "being physically constrained" result from something markedly different, i.e. the technological prediction of "likely next 'tokens'", which means that, if we ask DigiDan what it means for us, as humans, not to be free, the best it can do is to technologically predict "sequences of characters that often co-occur in written text". DigiDan cannot do what Dennett (and even the experts in his work) can do: first, experiencing us both aesthetically and epistemically (which means understanding from both our bodies and our words what our particular case is, i.e. what the particular meaning of our question is) and, second, being capable of original thinking (which means being capable of getting the fit between "rule" and "application", as we have seen: if our particular case is not included in his previous thinking, he can think in an original way and give us brilliant answers) – in other words, he, as a human being, is capable of getting the fit between his aesthetic and epistemic experience and reflection upon the meaning of not

being free (i.e. the "rule") and his aesthetic and epistemic experience and reflection upon the particular meaning of our question (i.e. the "application"), which is something that the caged parrot and DigiDan cannot do at all.

Thus, even though we may happen to judge the words "being physically constrained" as interesting, there is a sense in which we cannot trust them as reliable knowledge, in that they are nothing but the aesthetic shape of something that is far from being a true philosophical insight: again, they are underpinned by nothing but the technological prediction of "likely next 'tokens'", i.e. "sequences of characters that often co-occur in written text" – which means that, again, they are not, and cannot be, underpinned not only by living first-hand aesthetic and epistemic human existences but also by experiencing and reflecting upon other human existences both aesthetically and epistemically.

3. *Meta-sensemaking versus randomness (and why we trade the former for the latter)*

In other words, the caged parrot and DigiDan cannot make sense of what it means not to be free – more precisely, they are not capable of meta-sensemaking as the capability not only to aesthetically live human existences but also to epistemically reflect upon them and, finally, make sense of them, as we have seen.

In what follows, I will argue that the divergence between meta-sensemaking as what humans can do and what generative artificial intelligence can do correlates with the divergence between autonomy and randomness as what automation actually means.

The notion of autonomy is exceedingly complex from a philosophical perspective. In what follows, I will refer exclusively to the cornerstones of Kant's notion of autonomy, starting with its etymology, according to which being autonomous means "living under one's own laws"[4]. If we move from the etymological perspective to Kant's notion of autonomy, we learn that speaking of autonomy means speaking of "the property of the will by which it is a law to itself" (Kant 1785, 4, 440) as what opposes not only the notion of heteronomy but also, interestingly enough, the notion of "automaton" (Kant 1788, 5, 97) – interestingly enough, autonomy opposes the word from which the word "automation" itself derives, which means that autonomy opposes the word that essentially de-

[4] See LSJ at https://www.perseus.tufts.edu/hopper/morph?l=autonomos&la=greek#lexicon (accessed: May 2, 2024).

fines technology itself. In the first case, autonomy opposes heteronomy as living under others' laws, which Kant defines as nothing but "contingent" (Kant 1785, 4, 444). In the second case, autonomy opposes "automaton" as the case in which "the freedom of our will […] would at bottom be nothing better than the freedom of a turnspit, which, when once it is wound up, also accomplishes its movements of itself" (Kant 1788, 5, 97).

Also, from an etymological perspective, "automaton" means "accident"[5] and correlates with meaningful words: first, the verb "automatizo", which means to "act of oneself, act offhand or unadvisedly, […] to be done spontaneously or at random, […] haphazard, […] [to] introduce the agency of chance, […] [to] happen of themselves, casually"[6], second, the noun "automatismos", which means "that which happens of itself, chance"[7], and, third, the noun "Automatia", who is "the goddess of chance"[8].

Thus, the opposition between autonomy and automation is rather radical. In the case of autonomy, we address one of the essential notions on which Kant bases the notion of human being itself – thus, we may say that Dennett and the experts in his work, as humans, can be autonomous, which means that their answers to our questions can result from their autonomy. Conversely, in the case of automation, we address one of the essential notions on which the notion of technology itself is based (see also Rifkin 1995)[9] – thus, we may say that DigiDan, as technology, can be nothing but automated, which means that its answers to our questions can result from nothing but its automation. And, if it is true that speaking of automation means speaking of randomness, as we have seen, it is also true that DigiDan's answers to our questions can result from nothing but its randomness – there is a sense in which the definition of not being free as "not being physically constrained" can be nothing but random: resulting from nothing but "likely next 'tokens'", i.e. "sequences of characters that often co-occur in written text", is far from resulting from autonomous humans whose "will […] is a law to itself".

Thus, we may ask ourselves why we should trade Dennett and the experts in his work for DigiDan – and, more extensively, why

[5] See LSJ at https://www.perseus.tufts.edu/hopper/morph?l=automaton&la=greek#lexicon (accessed: May 2, 2024).

[6] See LSJ at https://www.perseus.tufts.edu/hopper/morph?l=automatizw&la=greek#lexicon (accessed: May 2, 2024).

[7] See LSJ at https://www.perseus.tufts.edu/hopper/morph?l=automatismos&la=greek#lexicon (accessed: May 2, 2024).

[8] See LSJ at http://stephanus.tlg.uci.edu/lsj/#eid=18224 (accessed: May 2, 2024).

[9] Especially from 1947 onwards, when the Ford Motor Company set up an automation department.

we should trade autonomy for automation[10]. Kant can continue to illuminate us. Whenever we trade autonomy for automation, which makes us as free as "the freedom of a turnspit", we can unburden ourselves from the burden of autonomy itself. Indeed, being autonomous means, first, that "moral law is possible" (Kant 1788, 5, 97), second, and consequently, that "imputation" (Kant 1788, 5, 97) is possible and, third, and consequently, that we are "culpable and deserving of punishment" (Kant 1788, 5, 100) whenever we (autonomously) fail. Conversely, being automated means, first, that "moral law" is replaced by a kind of "mechanism of nature" (Kant 1788, 5, 97) as nothing but a kind of "necessity of events in time" (Kant 1788, 5, 97), second, and consequently, that "imputation" is not possible and, third, and consequently, that we are not "culpable and deserving of punishment" whenever we fail – and, especially in our increasingly individualistic, competitive, accelerated and performing neoliberal societies, in which failure is almost unspeakable, unburdening ourselves from the increasingly unbearable burden of autonomy, which makes us "culpable and deserving of punishment" whenever we fail, may be precisely what we desperately try to do, sometimes consciously and sometimes unconsciously.

If we go back to the case of AI philosophers replacing human philosophers, we may think of DigiDan as an isolated case of testing. Yet, even though it is true that philosophy is markedly less characterised by the use of generative artificial intelligence than other areas typical of human creativity, it is also true that it may be a matter of time. On the one hand, the use of generative artificial intelligence exponentially affects professions from the least creative to the most creative, among which the various ways of writing are no exception, as several websites popularise[11]. On the other hand, the use of generative artificial intelligence exponentially affects students as future professionals who may naturalise it. As I myself can sense from my students as future creatives, if they are increasingly put under pressure by their increasingly individualistic, competitive, accelerated and performing educational environments, they may happen to increasingly use "design software [that] automates common processes. This allows you to save time. You can get designs […] quicker and enjoy a more efficient workflow. […] Its key feature is automation. […] [T]he software offers several tools for automating complicated processes. For example, it has

[10] I worked on the reasons why, from a philosophical perspective, we trade autonomy for automation in Chiodo 2022, 2023b and 2023c.

[11] For instance, see https://it.semrush.com/goodcontent/content-marketing-blog/ai-writing-tools/ (accessed: May 2, 2024).

an inbuilt library of stair and rail designs. […] You only need to enter a number into your variable to create the desired number of elements"[12]. Even in the case of the most creative (and, thus, pleasant) professions, if the pressure of overscheduling, overdoing and overperforming is taken to the extreme, technological automation, from totally standard "designs" to one's total alienation from one's (pleasant) creativity, may win out over the exercise of one's autonomy, i.e. autonomous creativity.

And, if it is true that even the most creative (and, thus, pleasant) professions are increasingly affected by the trade of human autonomy for technological automation, why should not it also happen to philosophers, who are increasingly asked to address the unprecedented complexity and uncertainty of the present time, from global health emergencies to global climate emergencies to global geopolitical crises to global economic crises to global crises of democracies and ideals to wars? More precisely, why should not it also happen to philosophers as humans who are exceedingly exposed to the increasingly unbearable burden of autonomously making sense of what increasingly seems to make no sense at all?

Yet, what AI philosophers lack, i.e. aesthetically living human existences, epistemically reflecting upon them and, finally, making sense of them, may also be what can save human philosophers, starting with aesthetically living human existences. Indeed, speaking of aesthetically living human existences means speaking not only of what underpins meta-sensemaking, as we have seen, but also of one of the most pleasant conditions of being human. Let us go back to our thought experiment according to which we struggle with the meaning of free will. Specifically, we are even faced with an existential crisis, tormented, in need of both making the right decision and living up to the exercise of free will and confused – which means precisely that we live aesthetically: any of our epistemic questions correlates with what we may define as our aesthetic questions, in that speaking of struggling, existential crisis, being tormented, being in need and being confused means speaking of sensations, emotions and feelings. And speaking of sensations, emotions and feelings means speaking of what can push us to give sensible answers even more than our epistemic reasons for answering our questions – our aesthetic reasons for answering our questions can be even stronger than our epistemic reasons. Whenever Dennett and the experts in his work, as humans, answer questions on free will, they can have both epistemic

[12] See https://academy.archistar.ai/top-ten-design-software-for-architects (accessed: May 2, 2024).

and aesthetic reasons for giving the most sensible answers they can, in that they know both what it means to be epistemically challenged and, especially, what it means to be aesthetically challenged by even torturing sensations, emotions and feelings. Once Dennett and the experts in his work, as humans, answer questions on free will, they can also make sense of their both epistemic and aesthetic lives and, thus, find a sense of resolution and even evolution, which works as an essential push to give the most sensible answers possible. Conversely, whenever DigiDan, as technology, answers questions on free will, it cannot have epistemic and, especially, aesthetic reasons for giving the most sensible answers it can, in that it does not know what it means to be epistemically challenged and, especially, what it means to be aesthetically challenged by even torturing sensations, emotions and feelings. Once DigiDan, as technology, answers questions on free will, it can do nothing but automatically "predict likely next 'tokens'" – and automatically "predict likely next 'tokens'" has nothing to do with making sense of both epistemic and aesthetic lives and, thus, finding a sense of resolution and even evolution: in the case of DigiDan, as technology, speaking of a sense of resolution and even evolution as an essential push to give the most sensible answers possible means speaking of nothing but a nonsense.

If my argument makes sense, we may try to consider at least the following two last remarks: the first is on a strictly philosophical issue and the second is on an issue between philosophy and technology policy.

First, we should stop thinking of technologies as kinds of anthropomorphic artefacts, which means not only that they cannot always be designed according to human prerogatives (see at least Sætra 2022), starting with human somatics, but also that they cannot always be used according to human perspectives, starting with human epistemic and, especially, aesthetic challenges. Conversely, we should intensify our philosophical work on the reasons why we are exponentially willing to trade the prerogatives of being human, i.e. autonomous, for the prerogatives of being technological, i.e. automated. If my argument makes sense, the starting hypothesis may be that what we are exponentially willing to trade is the unbearable burden of human autonomy for what can unburden us from it, i.e. technological automation – what we are exponentially willing to trade is human autonomy, as what is unbearably challenged by the unprecedented complexity and uncertainty of the present time, for technological automation, as what can unburden us from making sense of them, starting with our even torturing

sensations, emotions and feelings, i.e. our aesthetic lives, whenever we desperately fall into struggle, existential crisis, torment, need and confusion.

Second, work on technology policy should also be intensified. If it is true that AI philosophers cannot replace human philosophers, in that they lack essential conditions for giving us sensible answers, it is also true that we, as citizens, should always be put in a position to know what is the source of the answers to our questions. We cannot trust DigiDan (and the caged parrot) as a reliable source of knowledge, as we have seen. But, even though we decide to trust it for a variety of reasons, starting with the kind of desperate trade argued above, we should always have the right to know what is the source of the answers to our questions, in that it makes a big difference (in the future, if not in the present), as we have seen. Europe is well on its way to guarantee it[13], but more work is needed when it comes to education, starting with students exponentially affected by the use of generative artificial intelligence, who may become future professionals who may naturalise it, as we have seen. More precisely, when it comes to education, we, as (human) philosophers, may try to teach our students that meta-sensemaking may be far better than randomness even when it means desperately falling into struggle, existential crisis, torment, need and confusion, in that the former can give us what the latter cannot give us: making us not only epistemically live but also, and especially, aesthetically live – making us capable of sensations, emotions and feelings, and not atrophied, and, thus, making us capable of both the pleasure of feeling pleasure and the pleasure of feeling in itself even when we do not feel pleasure, but struggle, existential crisis, torment, need and confusion (from which, differently from DigiDan, we can learn essential lessons on what it means to be autonomous humans).

References

Anscomb, C., 2021, "Creative agency as executive agency. Grounding the artistic significance of automatic images". *The journal of aesthetics and art criticism* 79/4: 415-427.

Arielli, E., 2021, "Extended aesthetics. Art and artificial intelligence", *Proceedings of the European society for aesthetics* 13: 1-3.

Barale, Alice, ed., 2020, *Arte e intelligenza artificiale. Be My GAN*, Milano: Jaca Book.

[13] See https://artificialintelligenceact.eu/ (accessed: May 6, 2024).

Boden, M.A., 2010, "The Turing test and artistic creativity", *Kybernetes* 39/3: 409-413.
Boden, M.A., 2019, *From fingers to digits. An artificial aesthetic*, Cambridge: MIT Press.
Cetinic, E., She, J., 2022, "Understanding and creating art with AI. Review and outlook", *ACM transactions on multimedia computing, communications and applications* 18/2: 1-22.
Chiodo, S., 2022, "Human autonomy, technological automation (and reverse)", *AI & society* 37/1: 39-48.
Chiodo, S., 2023a, "Intelligenza artificiale e destino della forma", *Scienza e filosofia* 30/2: 210-234.
Chiodo, S., 2023b, *Technology and the overturning of human autonomy*, Cham: Springer.
Chiodo, S., 2023c, "Trading human autonomy for technological automation", in S. Lindgren, ed., *Handbook of critical studies of artificial intelligence*, Cheltenham-Northampton: Edward Elgar Publishers: 67-78.
Chiodo, S., 2024, *Predicted humans. Emerging technologies and the burden of sensemaking*, Abingdon-New York: Routledge-Taylor & Francis Group.
Coeckelbergh, M., 2017, "Can machines create art? *Philosophy & technology* 30/3: 285-303.
Elgammal, A., 2019, "AI is blurring the definition of artist. Advanced algorithms are using machine learning to create art autonomously", *American scientist* 107/1: 18-22.
Elgammal, A., Liu, B., Elhoseiny, M., Mazzone, M., 2017, "CAN. Creative adversarial networks, generating 'art' by learning about styles and deviating from style norms", *Arxiv*: 1706.07068.
Fazi, M.B., 2019, "Can a machine think (anything new)? Automation beyond simulation", *AI & society* 34: 813-824.
Goodfellow, I.J., 2016, "NIPS 2016 tutorial. Generative adversarial networks", *Arxiv*: 1701.00160.
Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014, "Generative adversarial nets", *Advances in neural information processing systems* 27: 2672-2680.
Kant, I., 1785, in Gregor, M.J., ed., *Groundwork of the metaphysics of morals*, Cambridge: Cambridge University Press.
Kant, I., 1788, in Gregor, M.J., ed., *Critique of practical reason*, Cambridge, Cambridge University Press.
Liddell, H.J., Scott, R., 1940, *A Greek-English lexicon* (LSJ), Oxford: Clarendon Press.

Manovich, L., 2019, Defining AI art: three proposals. AI and dialogue of cultures, Saint-Petersburg: Exhibition catalogue.
Mazzone, M., Elgammal, A., 2019, "Art, creativity, and the potential of artificial intelligence", *Arts* 8/1: 26.
Miller, A.I., 2019, *The artist in the machine. The world of AI-powered creativity*, Cambridge: MIT Press.
Rifkin, J., 1995, *The end of work. The decline of the global labour force and the dawn of the post-market era*, New York: Putnam.
Sætra, H.S., 2022, "Robotomorphy", *AI & ethics* 2: 5-13.
Schwitzgebel, E., Schwitzgebel, D., Strasser, A., 2023, "Creating a large language model of a philosopher", *Arxiv*: 2302.01339.
Steiner, S., 2017, "Art. Brought to you by creative machines". *Philosophy & technology* 30/3: 267-284.
Strasser, A., Crosby, M., Schwitzgebel, E., 2023, "How far can we get in creating a digital replica of a philosopher?", in R. Hakli, P. Makela, J. Seibt, eds., *Social robots in social institutions. Proceedings of "Robophilosophy 2022*", Amsterdam: OIS Press, 371-380.
Taylor, G.D., 2014, *When the machine made art. The troubled history of computer art*, New York: Bloomsbury.
Terzidis, K., Fabrocini, F., Lee, H., 2022, "Unintentional intentionality. Art and design in the age of artificial intelligence", *AI & society* 38: 1715-1724.
Turing, A.M., 1950, "Computing machinery and intelligence", *Mind* 59: 433-460.