Medij. istraž. (god. 30, br. 2) 2024. (31-50)
PREGLEDNI RAD
DOI: 10.22572/mi.30.2.2
Primljeno: lipanj, 2024.

A Critique of Anthropocentrism in the Evaluation(s) of Artificial Creativity

Primož Krašovec*

ABSTRACT

This article presents a critique of anthropocentrism in certain theoretical evaluation(s) of artificial creativity as a type of creativity exhibited by artificial intelligence (AI). By anthropocentrism, we mean the positing of human intelligence as a norm, in relation to which any different exercise in intellect can only be seen as a deviation and a sign of inferiority. We were able to identify several problems with regard to such an understanding of AI and its creativity: 1) it is unethical in the sense that it precludes any recognition of AI as other intelligence; 2) it stems from the misconception that human intelligence is fully transparent to us and that we have full access to the processes generating our own intelligence; 3) it results in a narrow and reductionist view of AI that misunderstands AI and overlooks most of its creative potential(s); 4) it is also incorrect with respect to human intelligence and creative processes.

The article consists of three main sections. The first one attempts to demystify a deceptively clear (self) understanding of human intelligence, which we analyse as a form of fetishism – since the only thing to which we have immediate access are the products of our intelligence, we tend to fetishise them due to our lack of access to the processes generating them. The second section is a critique of anthropocentrism as an attempt to use such a misleading impression of human intelligence as a norm for

*Dr. sc. Primož Krašovec, docent na Odjelu za sociologiju, Filozofski fakultet Sveučilišta u Ljubljani, Aškerčeva 2, 1000 Ljubljana, Slovenija, Telefon: +386 41 414 114, Email: primoz.krasovec@ff.uni-lj.si, ORCID: https://orcid.org/0000-0002-9013-0305

AI. The third section presents a more open and inclusive analysis of large language models (LLMs) that is also more appreciative of their creativity, as opposed to the common dismissive reactions they provoke in the fields of literature and humanities.

Keywords: anthropocentrism, creativity, artificial intelligence, fetishism, large language models

"Culture has constituted itself as a defence system against technics […]"
(Simondon, 2017: 9).

Introduction

Although presenting an immense technological breakthrough – a shift from heteronomous programming to autonomous learning (Alpaydin, 2021: 10-11), and thus from deterministic machines to machines that could work with randomness and uncertainty (Parisi, 2015) – deep learning (DL) AI did not raise much public concern when it was first introduced into mass everyday use in the 2010s. Or, to be more precise, the concerns raised by DL AI in the 2010s were not about AI as such, i.e. about the way it worked or the nature of its intelligence, but rather had to do with its purposes and (social, cultural and economic) effects. The first wave of DL AI criticism was largely focused on extraction (users were being harvested for data without adequate monetary compensation) and surveillance (user privacy was being violated), without paying much attention to DL AI's epistemic dimension. In other words, DL AI was still mostly perceived instrumentally, i.e. as a tool (of either shady tech companies or shady government agencies), while the criticism was directed at the way it was used and the effects of such use(s), such as privacy breaches (Zuboff, 2019), election manipulations (Aral and Eckles, 2019), increasing social and economic inequality (Eubanks, 2018), as well as the erosion of the democratic public sphere (Splichal, 2022).

Yet the introduction of generative DL AI also immediately triggered an avalanche of epistemological discussions on intelligence and artificial general intelligence (Tegmark, 2018). It is our contention that although earlier DL AI was already posing these epistemic questions objectively – i.e. by its very existence as a non-deterministic intelligent technology –, these did not come to the surface right away since *discriminative* DL AI could still be understood as mechanistic (meaning pre-programmed and deterministic). In other words, although it was, in fact, non-mechanistic (Hui, 2023) discriminative DL AI – i.e. AI that can recognise input such as faces and

fingerprints – could nevertheless be *perceived* as mechanistic since it still followed commands (while its inner modus operandi no longer followed heteronomous rules or programmes). From the user's point of view, discriminative DL AI is no different from a (mechanistic) washing machine, meaning that the difference between pressing a button to start a 40-degree colour cotton cycle and pressing a fingerprint scanner on a smartphone is negligent. Although necessitated by the fact that DL AI was developed to operate in situations in which mechanistic machines would be inadequate (Mendon-Plasek, 2021: 44), the technical revolution taking place was not plainly distinguishable *until* the introduction of *generative* DL AI. What makes generative DL AI so special is the fact that it not only works in a non-deterministic, non-mechanistic way, but also *shows* it. In contrast to discriminative DL AI, generative DL AI does not just recognise patterns in data; moreover, it generates new cultural content.

Even though the main underlying technical principle of generative DL AI is no different in its essence from previous discriminative AI algorithms that could, for example, recognise cats in photographs, the difference in terms of its cultural impact has been immense. In our interpretation, this cultural shift from a critique of the *uses* of discriminative AI to epistemic questions *about* generative AI and its intelligence – albeit the technical mode of operation remained the same – was caused by the change in the user experience of AI. Generative AI no longer takes orders, but reacts to prompts with unpredictable and sometimes erratic (Amoore, 2020: 108-129) results. In other words, its indeterminate mode of operation, previously concealed behind the user interface, immediately becomes apparent.

Additionally, generative AI uses language efficiently in a way that is not pre-programmed (talking to an LLM is a completely different user experience from, for example, talking to pre-programmed NPCs in computer games) – i.e. generative AI appears to be intelligent. Such intelligent behaviour exhibited by a machine is, in turn, confusing due to an ancient and entrenched prejudice that machines, although they might be faster, stronger and more capable than humans, only demonstrate this superiority in lowly technical tasks, and exclusively by working in a pre-programmed, mechanistic manner (Broussard, 2018). Accordingly, the immediate reactions to machines acting intelligently in ways and domains considered exclusively human (such as language, imagination, creativity) are either awe (machines have reached our level of intelligence …) or anxiety (… and nothing good can come from it) – or a combination of both.

In our view, the problematic common denominator of such reactions is anthropocentrism. They take human intelligence as the measure of all intelligence, and thus perceive other types of intelligence only in terms of their similarity to human in-

telligence – generative DL AI caused both hype and anxiety only to the extent that it seemed too similar to human intelligence when it was first introduced. Later on, the erstwhile alternating hype and anxiety were replaced by another cultural attitude towards generative DL AI: a dismissive one, which deflated both the pumped-up and panicked reactions, and tried to dissuade the users about generative AI being actually intelligent by reducing it to "mere statistics" (Crawford, 2021: 205). In this interpretation, generative AI is, despite appearances, not actually intelligent, and is in fact still just mechanistic, rote computation (Collins, 2018).

It is this currently prevalent attitude towards generative AI that will be the object of our critique. Our main thesis is that such an attitude is not only as anthropocentric as the hype and anxiety surrounding the phenomenon of AI (it still considers human intelligence as the norm and all other intelligence as a deviation), which is simultaneously an ethical issue in its own right since it precludes any possibility of the existence of a genuine *other*, non-human intelligence (as in all intelligence can only exist as an imitation of the human norm), but also that this version of anthropocentrism is based on an epistemologically questionable understanding of human intelligence, grounded in common misconceptions on how we think we think, rather than how we actually think.

Unknown Processes, Fetishised Results

One way to approach the current dismissive attitudes towards (generative) AI as a lesser, truncated version of the human norm would be to see them as a perversion of Turing's original idea on testing machine intelligence. Unlike the contemporary common-sense variations of what became known as the Turing Test, which presuppose that our own intelligence is transparent to us, and thus engender comparisons of AI to an imaginary (mis)understanding of human intelligence, Turing's (1950) original point was the exact opposite of the way in which the term is now used in discussions about AI, as well as in popular culture. Turing himself was not trying to establish whether machine intelligence equalled or could, in theory, equal human intelligence for the simple reason that human intelligence cannot be grasped immediately. Therefore, he changed the original question from "Can machines think?" (which is closer to today's meaning of the Turing Test) to the more appropriate "What happens when a machine takes part in an imitation game?" (Turing, 1950: 433-434). Consequently, the issue was no longer whether a machine thought exactly like a human, but whether it could "pass" as a thinking being by deceiving a human user that it is human. To repeat the crucial point: in Turing's view, there was no necessity for machine intelligence to work in the same way as human intelligence, i.e.

to imitate its operation. This imitation, which Turing established as a criterion for evaluating machine intelligence, happens in the next stage, in its performance (e.g. generating text in a natural language).

This difference is crucial since it eliminates any kind of anthropocentric criteria for evaluating machine intelligence from the very start – a machine is deemed intelligent if its linguistic performance equals human performance, regardless of the process leading to the said performance. But the reason that Turing changed the original question and devised the imitation game as a test for machine intelligence is not – at least not explicitly – a critique of anthropocentrism, but rather a result of another problem that continues to haunt all discussions about artificial and human intelligence alike. The problem in question is that human intelligence is not transparent even to humans themselves – accordingly, the whole anthropocentric edifice is problematic not only on ethical but also on epistemological grounds – since it is founded on a mistaken belief that direct (self)understanding of human intelligence via introspection is possible at all. Contrary to that belief, Turing (1950: 446-447) deemed introspection unreliable and solipsistic, arguing that the only way we can know that we are intelligent is by assessing our performance. Hence, Turing left the question of "real" human intelligence "black-boxed". We have no direct way of knowing how we achieve intelligent performance, we just know that we do – and since this is the only indicator of intelligence known to us, it makes sense to gauge other types of intelligence by how their performance measures against ours, without prejudice as to how they work "on the inside".

Turing's reluctance to probe both human and machine black boxes can be at least partially attributed to his personal need to keep his private sexual inclinations "black-boxed" and to pass undetected in the outside world (Bratton, 2015: 71-72); but the problem with AI (both now and then) is not – unlike male homosexuality in the early 20th century England – that its intelligence is stigmatised, and is therefore best kept hidden, but rather comes from the unwillingness on the part of culture to even acknowledge other intelligence(s), obstinately insisting that machine intelligence can only be an inferior version of the human norm. However, as already noted, this attitude rests on quite questionable assumptions about human intelligence.

In literature that criticises AI, much has been made about the Explainable AI (XAI) problem, but even the critiques of AI's opacity, however lucid otherwise, are still based on the premise that human intelligence is transparent and that we can grasp it immediately (Amoore, 2020: 1-25). In our view, the discussion on XAI is mistakenly framed as a discussion on ethics since its objective is to make AI fully transparent and thus amenable to direct human control – and there can be no ethical relation to either other humans or machines when both are completely programmed, i.e. ensla-

ved (Simondon, 2017: 141). When someone's or something's behaviour can be programmed, there is no need for ethics, whereas ethical relations can only be formed when we allow others to remain "black boxes", i.e. when we encounter (both living and machinic) autonomous, indeterminate beings. In this sense, our argument will run counter to the discussions on XAI, as we will not aim to make AI more transparent, but will instead try to show that human intelligence is less transparent than we imagine – that humans are black boxes as well.

We can start our assessment of the Explainable Human (XH) problem with a surprising glimpse into an actual human creative process and the stark way in which it differentiates from the common (self)understanding of human creativity:

> Walter Scott, that master of the 19th-century historical novel, wrote and revised and produced all his life, working industriously at novel after novel. Then towards the end of his life he had a series of debilitating strokes: he could, at the last, barely speak, and was almost unrecognisable as the man he had once been. And yet he kept writing. The very last few Waverley novels are fascinating: not 'good' by conventional metrics, but recognisably Scott in a free-associative sort of way, and extraordinary works: as if Scott was a writing machine that just continued churning out stories even after his conscious mind had been disengaged (Roberts, 2023).

What this example demonstrates is a stark contrast between the way in which we imagine we create and the actual processes of human (literary) creativity, which are usually hidden from our conscious (self)perception, but which can be brought to light in exceptional cases such as brain injury or mental illness. When tapping into our "regular" imagination, we might create by consciously applying our ideas and expressing our emotions, but in the example of post-stroke Scott, we can see that the actual process is much more out of our control and machine-like, with conscious control having a far lesser role in human creativity than we might have imagined, but also that there is an unconscious creative process that keeps on working even if conscious mental functions are severely impaired. We can thus separate the actual unconscious creative processes from our conscious awareness of their existence – these processes do not rely on us being aware of them, whereas being so oblivious to them, we also usually tend to overlook their contribution to human creativity, which seems transparent to us precisely because their crucial role remains hidden.

While covert creative processes do take place in our brains and minds, our subjective awareness has no access to them. Instead, what we do have access to are only the results of these processes (Bakker, 2018), meaning that they are the only available materials from which we can construct our common notions of how human creativi-

ty and intelligence in general work. Since we do not know what we do not know and we are not aware of what we are not aware of – what Bakker (2018) calls "neglect of neglect" (or "medial neglect) – we are prone to making up "just so" retroactive descriptions of the processes leading to creative results. In a movement surprisingly similar to commodity fetishism from the Marxian theory, "the process vanishes in its own result" (Heinrich, 2006), which in turn begins to appear as the origin of the process (Ranciere, 2016). To paraphrase Marx (1990: 166-167): We exercise our intelligence without being aware of it. And since we are, by the very design of our minds, precluded from access to the processes that lead to it, we fetishise human imagination, intuition and creativity, and at the same time tend to be dismissive towards forms of intelligence that diverge from these mystified appearances. Although "the observer does not have an exclusive, intimate access to the objects of its cognition and representation that would enable it to witness 'real' mental states" (Bach, 2009: 10), there still exists a false sense of intimacy with the results of these mental states and processes – we feel as though we, as sentient beings, should also be aware of our "real" mental processes, but since we are not, we literally make them up: "Because our minds aren't what they seem […] conscious introspection [is] misleading at best and complete fabrication at worst" (Pollack, 2014: 289).

In stark contrast, the development of AI implies designing intelligent *processes*, which thus has no option but to depart from fetishised results that appear in conscious introspection as a starting point of intelligence in order to recreate similar results with machinic processes. Alongside all commercial and other common uses of AI, its added epistemic value might be that AI, by encountering a host of theoretical and engineering problems and at the same time devising counter-intuitive ways of solving them (for example, by recasting the language generation process not as an application of grammatical rules, but as a stochastic process), reestablishes the importance of understanding *intelligence as a process*. Consequently, AI casts some doubt on our conceit regarding our understanding of our own intelligence and a lack of attention to the processes involved therein.

The design of our minds lets us experience only the results of our (unconscious) mental processes, but not the processes themselves. We therefore spontaneously develop something akin to an "antiprocess bias" by being fixated on the content of our awareness, such as feelings and ideas, and then fetishising this content. In our everyday experience of our own intelligence, the actual causation taking place is reversed: according to the way in which we experience our intelligence, we do not perceive unconscious emotional processes leading to the feelings; instead, we see those feelings as catalysts of creativity, while we grasp ideas as stepping stones towards ingenuity, etc. In short, results are experienced as causing (emotional and intellectual) processes, and not the other way around.

This anti-process bias necessarily involves the fetishism of human creativity and intuition, and its downside is a dismissive attitude towards AI as incapable of inventing anything new, as a tool merely mimicking and regurgitating the work of human authors (TAG, 2023). FKA Twigs's (2024) recent highly-publicised statement before the US Senate Judiciary Subcommittee on Intellectual Property incorporates both of those attitudes – the fetishisation of human experience combined with a belief that AI, since it is lacking human experience, is somehow inferior to human creativity, but at the same time frightening, since it can match its performance. Leaving aside the artist's quite justified concerns over unjust appropriation of artistic work via deepfakes, we can, as a way of exemplifying a characteristic contemporary cultural attitude towards generative AI, focus on her more existential concerns:

> […] my music, my dancing, my acting, the way that my body moves in front of a camera and the way that my voice resonates through a microphone is not by chance; they are essential reflections of who I am. My art is the canvas on which I paint my identity […] AI cannot replicate the depth of my life journey […] (Twigs, 2024: 1)

The main premise of the artist's reflection is that art is a direct expression of one's singular lived human experience: "Our creativity is the product of this lived experience […]" and "[…] the very essence of our being at its most human level […]" (Twigs, 2024: 2). Lived experience thus comes first, and is exclusively human, while artistic creativity expresses this lived experience. The problem with AI, from this perspective, is that it is not human and thus has no lived experience; AI creativity is accordingly judged to be fake (it has no relation to any lived experience) and is, at best, an inadequate imitation. But taking a less judgmental stance could allow for a problematisation of the usually assumed direct correlation between lived human experience and artistic creativity in the sense that perhaps the covert, unconscious processes of human creativity do not proceed from fetishised results, registered by awareness (experience), and as such might not be an exclusive privilege of the living. In other words, artistic creativity might not be the "essence of the [human] being" (Twigs, 2024: 1), but rather something inhuman exceeding ordinary human existence and its experience.

Creativity Without Emotions or Experience

To take another (by now already classic) example – in early 2023, Nick Cave (2023) joined the choir of AI creativity deniers on his blog, where, in a reply to a fan posting lyrics written by ChatGPT in the style of Nick Cave, he developed his own version of a common sense Turing Test, positing a naive (self)understanding of human creativity as the norm and AI creativity as a deviation, in a way that any *difference* between human and machine intelligence could only be framed as an inadequacy on the part of the machine; consequently, all conceivable machine intelligence would only be a pale imitation of human intelligence. As we described earlier, the problem lies in shifting the criteria of machine intelligence from the performance (like in Turing's original idea) of intelligence to the process behind it. A fair share of criticism of AI creativity coming from the world of arts and humanities gets stuck in the loop of first: the simulative paradigm (Fazi, 2019), defined by the inability to even imagine an intelligence that would not be bound to the human norm, and second: the naive common-sense view of this norm itself.

With regard to Cave (2023):

> What ChatGPT is, in this instance, is replication as travesty. ChatGPT […] cannot create a genuine song. It could perhaps in time create a song that is, on the surface, indistinguishable from an original, but it will always be a replication, a kind of burlesque.
>
> Songs arise out of suffering, by which I mean they are predicated upon the complex, internal hu-man struggle of creation and, well, as far as I know, algorithms don't feel. Data doesn't suffer. ChatGPT has no inner being, it has been nowhere, it has endured nothing, it has not had the audacity to reach beyond its limitations, and hence it doesn't have the capacity for a shared transcendent experience, as it has no limitations from which to transcend. ChatGPT's melancholy role is that it is destined to imitate and can never have an authentic human experience, no matter how devalued and inconsequential the human experience may in time become.
>
> […] Writing a good song is not mimicry, or replication, or pastiche, it is the opposite. It is […] part of the authentic creative struggle that precedes the invention of a unique lyric of actual value; it is the breathless confrontation with one's vulnerability, one's perilousness, one's smallness, pit-ted against a sense of sudden shocking discovery; it is the redemptive artistic act that stirs the heart of the listener, where the listener recognises in the inner workings of the song their own blood, their own struggle, their own suffering. This is what we humble humans can offer, that AI can only mimic […] this song is bullshit, a grotesque mockery of what it is to be human […]

Cave never bothers to compare the quality of his own song to the song written by ChatGPT at the level of linguistic performance, but goes straight to the process – AI has no "authentic human experience" and has not endured emotional "suffering." Since it has no emotions or experience of its own, it is therefore "destined to imitate" human culture and artistic expression and will forever remain "mimicry, replication, pastiche" – an argument that draws dramatic conclusions, but is, in its structure, essentially the same as FKA Twigs's. The incorrect notion and the object of our critique in both cases is first: the assumption that there is a direct causal connection between human emotional experience and the process of human creativity, and second: that human creativity (or intelligence in general) is the only conceivable creativity (or intelligence), such that anything different from human creativity can only be its forever inadequate replication.

As documented by Sautoy (2019: 189-190), music critics routinely evaluate AI generated music positively when agnostic as to its source, but as soulless and lacking emotion when they know that it was generated by AI. The logic involved is characteristic of what we call the antiprocess bias, whereby results – human emotions – are fetishised as sources of creativity. Since AI has no emotional experience, in the eyes of its critics, it cannot generate a genuine creative process. Any notion of a creative process that could be generated without an emotional experience is dismissed out of hand. Generative AI is consequently reduced to a low-quality imitation of human creativity, with the latter constituted as the norm. To quote the influential technology critic James Bridle (2023): in "AI creativity" [*in quotation marks in the original*] "there is no true originality […] only very skilled imitation and pastiche […]." As a result, AI creativity is dismissed as second-rate and worthless.

While the common understanding of human creativity within the fields of arts and humanities could be summed up with Trkaj's (2004) verse that poetry is made of "genuine emotions that are put on paper", Rastko Močnik (2006) presented the opposite view in his seminal study of Prešeren's poetry *Julija Primic v slovenski književni vedi* ["Julija Primic in Slovenian Literary Studies"]. In the versions of Prešeren's life's work taught in Slovenian schools, Julija Primic is usually featured as the poet's muse, and his feelings for her are, by extension, the source of his creativity. But this interpretation, although in line with artists' common sense and their understanding of human creativity, cannot answer the crucial question regarding Prešeren's creativity – how can something as extraordinary as Prešeren's immaculate sonnets, written in the Slovenian language that was thought to be low-brow and therefore inadequate for higher literary pursuits at the time (the early 19th century), originate from something as ordinary as a crush on a younger and unattainable (due to class distinctions) woman during a midlife crisis? In other words, if poetry was

really a direct translation of emotions into words, any middle-aged creep stalking a younger woman could write *Sonetni venec* ["A Wreath of Sonnets"] – but that is hardly the case.

As shown by Močnik (2006), even human poetry is not an expression of personal emotions related to lived experience. Any work of poetry establishes connections with and proceeds from other works of poetry, and in Prešeren's case, his relation to Petrarch was much more significant for his poetry than his feelings for Julija – if these feelings even existed at all, and Julija was not just a prop, since unrequited love is a standard trope in sonnet writing. On the other hand, "poetry" as an expression of genuine emotions certainly can appear in many private love letters and high school love notes, but is hardly on the level of the likes of Prešeren, Cave or Twigs. Poetry is much less an expression of our personal emotions, and much more 'algorithmic' than we are willing to imagine; however, since once again, the real processes behind creativity are unconscious, we tend to fetishise whatever is accessible to us (in this case, emotions and experience). Algorithms of poetry are not explicit and do not constitute unambiguous rules for poets to follow – instead, the creativity of poets works as cultivated intuition (Pedwell, 2023). The structure of poetry only becomes intelligible retroactively, through literary analysis, and is not something apparent in advance. In short, in order for poetry to be poetry, it has to be created intuitively – and it is precisely because it is created intuitively that it has to involve the misrecognition of the creative process behind it.

If creative human writing is not an expression of emotions, what about the other element so insisted upon by Cave and Twigs as a singular determinant of human creativity: one's lived experience? Deleuze and Guattari captured something similar to Močnik's theory when they remarked that

> [w]e dwell on the art of the novel because it is the source of a misunderstanding: many people think that novels can be created with our perceptions and affections, our memories and archives, our travels and fantasies, our children and parents, with the interesting characters we have met and above all, the interesting character who is inevitably oneself (who isn't interesting?), and finally, with our opinions holding it all together (Deleuze and Guattari, 1994: 170).

One's lived experience can certainly be used as writing material, but what distinguishes ordinary experience-sharing newspaper columns from auto-fiction on the level of, for example, Elena Ferrante, is that in the second case, the experience is alienated, i.e. no longer treated as something personal, but precisely as material to be mediated by literary form(s). In this sense, while authors' lived experiences can

become their writing material, they are treated no differently than completely made up lives and experiences – once again, as in the case of emotions, we see that literary forms and intuitive writing algorithms, not personal experience, are decisive.

In other words, not only does the creative human process not rely on first-hand experience, but it actually involves keeping a distance from it in order to turn this experience into writing material. And since it is possible to write about anything, even (as in the case of science fiction) completely made up other or future worlds, personal experience is equally irrelevant for human creativity as personal emotions. Obstinate opinions on AI as not being "truly" creative – since it has no emotions or experience – reveal more about our prejudice towards machines than about the actual creative processes taking place within either humans or machines.

If human creativity was indeed based on experience, it would be much more diminished than it actually is since the whole point of imagination is to *imagine* (and write about) something one hasn't experienced oneself. It is precisely the impersonality of literature that allows us to escape the narrow confines of our own lives and lived experiences and be creative – human imagination and creativity represent the very act of transcending our experience. The same can be said about human language in general – it is a cognitive technology that, by being symbolic, allows us to dissociate from our immediate lived situations and imagine other situations (Leroi-Gourhan, 1993: 178-216).

In his discussion on the problem of understanding in LLMs, Arcas (2022: 188) puts forth a telling example of Helen Keller, a deaf and blind writer, who could still imagine sounds and colours even though she had no direct, lived experience of them – and probably none of the AI critics would deny her the capacity for creativity in the same way that it is routinely denied to AI. Arcas's main point is worth quoting in full:

> […] the socially learned aspect of perception is likely more powerful than many of us realise; shorn of language, our experiences of many sensory percepts would be far less rich and distinct. In fact, there are many nuances we are perfectly capable of perceiving but are blind and deaf to in precisely the ways Keller was not: our deficit is in language and culture, not in sensory organs (Arcas, 2022: 189).

It is precisely language (and writing algorithms) that in many ways shapes our perception(s) and experience(s), and there are very few instances (if any at all) of "pure" human experience, which is unmediated by language and just waiting for creative "expression". Therefore, what makes human creativity (and, in large part,

also human experience itself!) possible are language algorithms, which can, as in the case of Keller, compensate for the lack of lived experience. The same can then be said of machine creativity – if the lack of immediate, lived experience of the world does not prevent creativity in humans, machines can also, at least in theory, be creative.

Language as Its Own World

Since they are not living beings, it is only logical that LLMs have no lived experience, much less a subjective, direct and self-aware one, which is characteristic of humans. Moreover, LLMs have no access and exposure to the world outside language, meaning that their "experience" is not only radically different from human experience (Fazi, 2018: 25-30), but also limited to the world of language as such. In short, not only do LLMs have no inner, emotional and mental world to creatively express, but they also have no relation to the outside world since their world consists entirely of immense volumes of text.

Be that as it may, as we have established earlier, experience is not a necessary prerequisite for creativity and can only serve the creative process once it has been sufficiently alienated. Admonishing LLMs for a lack of human-like understanding of language is a staple of AI criticism (Bogost, 2022), and is usually related to AI's lack of experience of the outside world and the accompanying emotional valence ascribed to that experience (Skubic, 2024). While it is true that LLMs' "understanding" is radically different from human understanding, grounding even human understanding in immediate lived experience is problematic since immediate experience can act as an obstacle to understanding, in the same way it acts as an obstacle to creativity. Science is similar to literature in the sense that it also goes beyond immediate experience, abandoning this experience so as to develop superior understanding, just like literature abandons it in order to develop superior creativity. In distinction to the anthropocentric positing of human experience as the norm of "true" understanding, Amoore (2020: 29-55) points out some remarkable similarities in the ways in which both science and AI bypass experience in order to reach that, which either cannot be experienced or is hard to understand due to the false certainty generated by experience. In this sense, AI has the advantage of having no experience to alienate, overcome or bypass.

If we suspend anthropocentric prejudice, "large language models are the first to illustrate the way in which language understanding and intelligence can be dissociated from all the embodied and emotional characteristics we share with each other and with many other animals" (Arcas, 2022: 194). While "constraining machines

to retrace our steps – or the steps of any other organism – would squander AI's true potential: leaping to strange new regions and exploiting dimensions of intelligence unavailable to other beings," (Browning, 2020) the astonishing thing about LLMs is precisely how they can develop new, alien modes of understanding, and use language without any of the features previously thought to be indispensable for language use, understanding and creativity. There is an immense amount of creativity already inherent in the very way LLMs function – in the sense that they neither contain any pre-programed rules for language generation nor do they present an application of any human-made theory of language (Sejnowski, 2018: 245-260). Before generating any new text, LLMs already generate a whole new perception and understanding of language, based on a hyper-dimensional closeness and distance between words rather than on meaning in an ordinary human sense.

Considering this, the standard critique of LLMs – that they do not understand language – loses much of its impact since the understanding that LLMs lack is actually human understanding based on the experiential recognition of meaning. However, as shown by Fazi (2021: 63), AI is not so wanting in its lack of human understanding, as it demonstrates that human understanding is only one of the possible means for using language. AI involves no human meaning or understanding, which is by design rather than omission (Halpern, 2014: 207). When generating a new text, LLMs show what language can do when it is released from the constraints of human meaning and understanding. What is, in human creativity, a singular achievement of dispensing with one's all-too-human experience(s) and playing with the very limits of language, is a given for AI: "[…] artificial intelligence need not – and should not – be confined to simply imitating human intelligence" (Browning, 2020).

Similar to how common (mis)understandings of human creativity act as an obstacle to the appreciation of artificial creativity, those same (mis)understandings of the human use of language and the role of understanding in it act as a major obstacle to the appreciation of LLMs. A common critique of LLMs is that they use language superficially, without understanding, whereby understanding is regarded as a connection between language and the world, and it consequently entails, for example, connecting the word "tree" with the tree as an object. By learning exclusively from text and having no perception of the outside world, LLMs are in this sense truly Derridean, as for them, there is nothing outside the text (Derrida, 1998: 158) – LLMs do not and cannot understand language by relating words to objects. But the question is: do humans?

According to Ducrot's seminal theory of pragmatics, "words do not mean anything" (Ducrot, 2009: 13), implying that, in most cases, words do not refer to or only indirectly refer to things outside of social situations, mediated by language. When

we say (Ducrot's favourite example) "the weather is nice", it is not an observation about the world communicated to another person, but rather an argument meant to convince the other person to join us for a walk (Ducrot, 2009: 48-55). Situations of using language to describe the world are rare and exceptional, whereas we mostly use language to influence others. For example, when we stress that "it is almost eight", we are not merely stating a fact about the world; instead, we want to urge our company to hurry up (Ducrot, 2009: 55-60). In this sense, the actual human use of language is already somewhat detached from the world and is limited to its social dimension. LLMs take this detachment of language from the world even further:

> The algorithm adapts to the world of text itself – the statistically relevant ways humans deploy symbols – not to the world as such. It does not occupy our niche, the niche of social beings who use language for diverse reasons, but its own: the regular interplay of signs. […] language might form its own ecosystem, into which a being might fit without human-like intelligence (Browning, 2020).

In the case of LLMs, language forms its own world and, as such, not only allows, but also calls for other-than-human forms of intelligence. While natural human language is already detached from the world by design, i.e. by being abstract, symbolic and part of a specifically human external technics (Leroi-Gourhan, 1993: 107), LLMs further uphold this logic by also removing language from the social world and treating it as their own world. In this situation, the only thing that matters are relations between words, and deep learning AI is exceptionally good at recognising these relations. What LLMs generate is pure, self-referential language that is neither a description of the world nor an expression of sensations or emotions.

At the same time, language as used and generated by LMMs is also not a "symbolic structure" composed of pure syntactic and grammar rules. Instead, LMMs treat language as a context-sensitive practical skill without regard for its formal rules as imagined by linguistics (Browning and LeCun, 2022). They combine words according to patterns, recognised in their practise, but this does not amount to LLMs being "stochastic parrots" (Bender et al, 2021), the latter implying their unintelligent and uninventive mimicry of human (linguistic) intelligence. Preferably, it might be more productive to understand LLMs as "stochastic chameleons" (Milliere, 2022), i.e. not as an inadequate intelligence in comparison to the human norm, but as other, alien intelligence, masquerading as a human one.

Regarding the question of linguistic creativity of such an alien machine intelligence, common criticism sees LLMs as unoriginal in the sense that they merely repeat what they learn in practise, meaning that they cannot be truly creative. While it is

true that LLMs cannot generate writing from their own emotions and experience (but this is in itself, as we have shown earlier, a rather naive understanding of human creativity), we can, if we adopt a more open-minded and appreciative attitude towards AI, still discern traces of genuine creativity in the way(s) they combine and (re)arrange text. Considering this, Marche (2023) offers an interesting insight: LLM text composition is very similar in form to the way DJs in early hip hop culture assembled elements of classic funk, soul and R&B songs in new combinations without "really" composing anything new. These combinations omitted choruses while emphasising the previously under-appreciated "breaks", thus revolutionising the popular culture of the day. Generative AI can be seen as a similar "archive creativity", whose originality lies not so much in new content, but in the surprising and novel ways it mixes existing cultural content – and this type of creativity is made possible precisely because DL AI works differently than human intelligence. In this interpretation, the difference between the two – contrary to common criticism that sees this difference as a sign of the inadequacy of AI compared to the human "original" – is actually the source of creativity of LLMs, which may be viewed as "fiction machines" (Bottou and Schölkopf, 2023).

Conclusion

Our theoretical investigation of the common critiques of artificial creativity has shown that they are, in most cases, misguided since they first: posit human creativity as the norm every other creativity is supposed to adhere to, and second: are based on a naive and illusory conception of human creativity. Human creativity itself is primarily an unconscious process to which we have no direct access, leading us to fetishise its results and reverse its causality – to the extent that we understand our creativity as proceeding from emotions and experiences, and then set these as norms for all other types of creativity to follow. Within that approach, artificial creativity inevitably comes off as a second-rate imitation.

If we, however, adopt a more open-minded and accepting attitude towards machine intelligence and creativity, we can see that other creative processes are possible and that human creativity, although undoubtedly very special and precious in its own right, is in no way the only possible creativity – especially when it comes to language and writing since they are, which also applies to their use among humans, precisely the technologies of abstraction that go beyond personal experience and allow for imagination, which is the very precondition of creativity. This means that emotions and experience are, even in humans, largely irrelevant when it comes to creativity, whereas an encounter with machine creativity can deepen our understanding of creativity in its human form.

Acknowledgment

The research leading to this article was carried out with the support of the research programme group known as The Issues of Autonomy and Identity in the Era of Globalisation (P6-0194), which is subsidised by ARIS (Slovenian Research and Innovation Agency).

LITERATURA

Alpaydin, E. (2021) Machine Learning. Cambridge, MA: The MIT Press. ISBN 9780262365369.

Amoore, L. (2020) Cloud Ethics. Durham, NC: Duke University Press. ISBN 978-1-4780-0831-6.

Aral, S. and Eckles, D. (2019) "Protecting Elections from Social Media Manipulation", Science, 365 (6456), 858-861. Doi: 10.1126/science.aaw8243.

Arcas, B. (2022) "Do Large Language Models Understand Us?", Daedalus, 151 (2), 183–197. Doi: https://doi.org/10.1162/daed_a_01909.

Bach, J. (2009) Principles of Synthetic Intelligence. Oxford: Oxford University Press. ISBN 978-0195370676.

Bakker, S. (2018) "Exploding the Manifest and Scientific Images of Man", Three Pound Brain. https://rsbakker.wordpress.com/2019/09/27/exploding-the-manifest-and-scientific-images-of-man-2/, downloaded on November 28, 2024.

Bender, E., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021) "On the Dangers of Stochastic Parrots", FAccT '21, pp. 610-623. Doi: https://doi.org/10.1145/3442188.3445922.

Bogost, I. (2022) "ChatGPT is Dumber Than You Think", The Atlantic. https://www.theatlantic.com/technology/archive/2022/12/chatgpt-openai-artificial-intelligence-writing-ethics/672386/, downloaded on November 28, 2024.

Bottou, L. and Schölkopf, B. (2023) "Borges and AI", AXiv. https://arxiv.org/abs/2310.01425v2, downloaded on November 28, 2024.

Bratton, B. (2015) "Outing Artificial Intelligence", pp. 69-80. In: M. Pasquinelli: Lüneburg: Me-son. ISBN 978-3-95796-066-5.

Bridle, J. (2023) "The Stupidity of AI", The Guardian. https://www.theguardian.com/technology/2023/mar/16/the-stupidity-of-ai-artificial-intelligence-dall-e-chatgpt, downloaded on November 28, 2024.

Broussard, M. (2018) Artificial Unintelligence. Cambridge, MA: The MIT Press. ISBN 9780262537018.

Browning, J. (2020) "Learning Without Thinking", Noema. https://www.noemamag.com/learning-without-thinking/.

Browning, J. and LeCun, Y. (2022) "AI and the Limits of Language", Noema. https://www.noemamag.com/ai-and-the-limits-of-language/, downloaded on November 28, 2024.

Cave, N. (2023) "I Asked ChatGPT to Write a Song in the Style of Nick Cave and This Is What It Produced", The Red Hand Files. https://www.theredhandfiles.com/chat-gpt-what-do-you-think/, downloaded on November 28, 2024.

Crawford, K. (2021) Atlas of AI. New Haven, CT: Yale University Press. ISBN 9780300252392.

Deleuze, G. and Guattari, F. (1994) What is Philosophy? New York: Columbia University Press. ISBN 9780231079891.

Derrida, J. (1998) Of Grammatology. Baltimore, MA: Johns Hopkins University Press. ISBN 0-8018-1879-6.

Ducrot, O. (2009) Slovenian lectures. Ljubljana: Pedagoški inštitut. ISBN 978-961-270-014-0.

Eubanks, V. (2018) Automating Inequality. New York: St. Martin's. ISBN 9781250074317.

Fazi, B. (2018) Contingent Computation. Lanham, MD: Rowman & Littlefield. ISBN 978-1-78660-609-9.

Fazi, B. (2019) "Can a Machine Think (Anything New)?", 34, pp. 813-824. Doi: https://doi.org/10.1007/s00146-018-0821-0.

Fazi, B. (2021) "Beyond Human", Theory, Culture & Society 38 (7-8), 55-77. Doi: https://doi.org/10.1177/0263276420966386.

Halpern, O. (2014) Beautiful Data. Durham, NC: Duke University Press. ISBN 978-0-8223-7632-3.

Heinrich, M. (2006) "A Thing with Transcendental Qualities", Monthly Review. https://mronline.org/2006/11/03/a-thing-with-transcendental-qualities-money-as-a-social-relationship-in-capitalism/, downloaded on November 28, 2024.

Hui, Y. (2023) "ChatGPT", E-flux. https://www.e-flux.com/journal/137/544816/chatgpt-or-the-eschatology-of-machines/, downloaded on November 28, 2024.

Marche, S. (2023) "The Future of Writing Is a Lot Like Hip-Hop", The Atlantic. https://www.theatlantic.com/technology/archive/2023/05/generative-ai-novel-writing-experiment-stephen-marche/673997/, downloaded on November 28, 2024.

Marx, K. (1990) Capital I. London: Penguin. ISBN 978-0140445688.

Leroi-Gourhan, A. (1993) Gesture and Speech. Cambridge, MA: The MIT press. ISBN 0-262-12173-5.

Mendon-Plasek, A. (2021) "Mechanized Significance and Machine Learning", pp. 31-78. In: Rob-erge, J. and Castelle, M.: The Cultural Life of Machine Learning. New York: Palgrave Macmillan. ISBN 978-3-030-56288-5.

Milliere, R. (2022) "Moving Beyond Mimicry in Artificial Intelligence", Nautilus. https://nautil.us/moving-beyond-mimicry-in-artificial-intelligence-238504/, downloaded on November 28, 2024.

Močnik, R. (2006) Julija Primic v slovenski književni vedi. Ljubljana: Sophia. ISBN 978-961-6294-82-9.

Parisi, L. (2015) "Instrumental Reason, Algorithmic Capitalism, and the Incomputable", pp. 125-137. In: M. Pasquinelli: Lüneburg, Meson. ISBN 978-3-95796-066-5.

Pedwell, C. (2023) "Intuition as a Trained Thing", Subjectivity 30, 348-372. Doi: https://doi.org/10.1057/s41286-023-00170-x.

Pollack, J. (2014) "Mindless Intelligence", pp. 279-293. In: Vargas, P., Di Apolo, E., Harvey, I. and Husbands, P.: The horizons of evolutionary robotics. Cambridge, MA: The MIT Press. ISBN 9780262026765.

Ranciere, J. (2016) "The Concept of Critique and the Critique of Political Economy", pp. 86-236. In: Althusser, L., Balibar, E., Establet, R., Macherey, P. and Ranciere, J.: Reading Capital London: Verso. ISBN 9781784781415.

Roberts, A. (2023) "It Is a Beast That Needs to Be Tamed", The Guardian. https://www.theguardian.com/books/2023/nov/11/chatgpt-needs-to-be-tamed-bernardine-evaristo-jeanette-winterson-put-ai-to-the-test, downloaded on November 28.

Sautoy, M. (2019). The Creativity Code. Glasgow: 4th estate. ISBN 9780674240407.

Sejnowski, T. (2018) The Deep Learning Revolution. Cambridge, MA: The MIT Press. ISBN 9780262038034.

Simondon, G. (2017) On the Mode of Existence of Technical Objects. Minneapolis: Univocal. ISBN 9781517904876.

Skubic, A. (2024) "Umetna inteligenca se ne more kesati namesto nas", Airbeletrina. https://airbeletrina.si/umetna-inteligenca-se-ne-more-kesati-namesto-nas/, downloaded on November 28.

Splichal, S. (2022) Datafication of Public Opinion and the Public Sphere. London: Anthem. ISBN 9781839984501.

Tegmark, M. (2018) Life 3.0. London: Penguin. ISBN 978-1101946596.

Turing, A. (1950) "Computing Machinery and Intelligence", Mind 59 (236), 433-460. Doi: https://doi.org/10.1093/mind/LIX.236.433.

TAG (2023) "The Authors Guild, John Grisham, Jodi Picoult, David Baldacci, George R.R. Martin, and 13 Other Authors File Class-Action Suit Against OpenAI", The Authors Guild. https://authorsguild.org/news/ag-and-authors-file-class-action-suit-against-openai/.

Trkaj, R. (2004) "Moja muska". In: Moja muska. Ljubljana: NIKA records. downloaded on November 28

Twigs, F. (2024) "Statement for US Senate Judiciary Subcommittee on Intellectual Property", US Senate. https://www.judiciary.senate.gov/imo/media/doc/2024-04-30_-_testimony_-_twigs.pdf, downloaded on November 28.

Zuboff, S. (2019) The Age of Surveillance Capitalism. London: Profile. ISBN 978-1610395694.

Medij. istraž. (god. 30, br. 2) 2024. (31-50)

Kritika antropocentrizma u procjeni umjetne kreativnosti

Primož Krašovec

SAŽETAK

Ovaj članak predstavlja kritiku antropocentrizma u određenim teorijskim ocjenama umjetne kreativnosti kao kreativnosti koju iskazuje umjetna inteligencija (UI). Pojam antropocentrizma podrazumijeva postavljanje ljudske inteligencije kao norme, u odnosu na koju se svaka varijacija može promatrati isključivo kao odstupanje i znak inferiornosti. Iz takva poimanja umjetne inteligencije i njezine kreativnosti proizlazi nekoliko problema: 1) neetično je s obzirom na to da u korijenu zatire svako razumijevanje umjetne inteligencije kao drugog oblika inteligencije; 2) polazi od pogrešne pretpostavke da nam je ljudska inteligencija sasvim dokučiva i da uživamo puni pristup procesima iz kojih proizlazi naša vlastita inteligencija; 3) za posljedicu ima uskogrudan pogled na umjetnu inteligenciju koji podrazumijeva pogrešno shvaćanje i previđanje većine njezinih kreativnih potencijala; 4) također njeguje zablude u vezi s ljudskom inteligencijom i kreativnim procesima.

Članak se sastoji od triju glavnih dijelova. U prvome se nastoji demistificirati naizgled jasno (samo)razumijevanje ljudske inteligencije, koja se ovdje analizira kao oblik fetišizma: s obzirom na to da imamo neposredan pristup isključivo ishodima svoje vlastite inteligencije, skloni smo ih fetišizirati u nedostatku pristupa procesima iz kojih proizlaze. Drugi se dio svodi na kritiku antropocentrizma kao nastojanja postavljanja takve pogrešne slike o ljudskoj inteligenciji kao norme za UI. U trećem dijelu iznosi se otvorenija i inkluzivnija analiza velikih jezičnih modela (VJM), koja pridaje veći značaj njihovoj kreativnosti od uobičajenih dismisivnih reakcija na kakve se nailazi u području književnosti i humanističkih znanosti.

Ključne riječi: antropocentrizam, kretaivnost, umjetna inteligencija, fetišizam, veliki jezični modeli