CAN ARTIFICIAL INTELLIGENCE (RE)DEFINE CREATIVITY?

Philosophical, Ethical and Legal Aspects

Dessislava Fessenko
Research Fellow, EthicAI=LABS Project, Goethe Institute Southeast Europe
dessislava.fessenko@fessenko.com

Abstract: What is the essential ingredient of creativity that only humans – and not machines – possess? Can artificial intelligence help refine the notion of creativity by reference to that essential ingredient? How / do we need to redefine our conceptual and legal frameworks for rewarding creativity because of this new qualifying – actually creatively significant – factor?

Those are the questions tacked in this essay. The author's conclusion is that consciousness, experiential states (such as a raw feel of what is like to be creating) and propositional attitudes (such as intention to instigate change by creating) appear pivotal to qualifying an exploratory effort as creativity. Artificial intelligence systems would supposedly be capable of creativity if they could exhibit such states, which philosophers and computer scientists posit as conceptually admissible and practically possible.

The existing legal framework rewards creative endeavours by reference to the novelty or originality of the end result. But this bar is not insurmountable for artificial intelligence. Technically speaking, artificial intelligence systems can create works that are novel and/or original. Are we then prepared to grant to those systems the legal status of "creators" in their own right? Whom should the associated benefits and rewards be assigned to? How does the position change (or not) based on the qualifying factors set out above? Should – and if, how – the general public benefit from inventions / creative works of artificial intelligence systems if troves of personal data are the key component that fueled and informed creative choices?

1 Introduction

Creativity is considered a sacrosanct realm of human beings. Purportedly, only they can possess and project in the material world the capability and skill to originate something new, original and aesthetically satisfactory. But is this unconditionally and irrevocably true? May the tables turn if novelty, originality and aesthetic judgement come about as a result of the deployment and operations of an artificial intelligence ("**AI**") system? And what if that AI system churns out a result irrespective of the initial instructions or input provided by the humans involved in the process? How about if AI can and acts autonomously? How then do these outcomes change our understanding and the notion of creativity?

Those are the questions that this essay purports to address. To provide definitive answers would probably be a tall order for the author. To entertain and contemplate (speculative) scenarios would be part of the

journey. And, ultimately, to put forward possible propositions for all of us to ponder upon is the humble mission that the author wishes to embark upon.

Along the way, we will inquire and contemplate a range of instrumental questions. In particular, what creativity is (Chapter 2). whether AI can be creative (Chapter 3), whether AI can refine the notion of creativity by reference to its essential ingredient (Chapter 4) and if and how we need to revisit our conceptual and legal frameworks for rewarding creativity so that they reflect the morally and socially significant element of creativity (Chapter 5).

2 Creativity, unpacked

Philosophy and psychology have gone to some lengths exploring the concept of creativity. Thinkers (from Plato and Aristotle to Berys Gaut) and scientists (from Henri Poincare to Marcus du Sautoy) have dwelled upon the essence, dimensions and variations of creativity and tested the limits of the various concepts thereof. For an extensive overview of the various schools of thought, see Gaut's "*The Philosophy of Creativity*".[1] I will focus here on the theories of most relevance to the topic of this essay.

2.1 Concepts of creativity: mental capacity or experiential mental process

1.1.1 The Computational Theory of Creativity

An influential account of creativity is presented by Margaret Boden's computational theory of creativity.[2] On Boden's view, human creativity is a mental capacity to generate "*ideas that are new, surprising and valuable*".[3] It is not to be regarded as a "*faculty*" (such as the various senses) but as an "*aspect of human intelligence*". This mental capacity is founded on the deployment of a wide range of physiological abilities such as noticing, perception, memory, contextualising, associative powers and recognising analogies, conceptual thinking, and reflective self-thinking.[4] It allows for generation of ideas by exploring, bending or breaking off, and, thus, transforming established conceptual spaces delineated by constraints (such as preconceptions or rules).[5]

These generative processes amount to creativity in a particular instance only if the ideas produced are new, surprising and of value, in Boden's view. They constitute the so-called "radical creativity" — the outstanding form typically of interest — as, in Boden's view, everyone is capable of some form/degree of creativity in daily life. Yet, the radical manifestations thereof merit recognition and reward.

Boden reckons an idea as *new* if it is such at least to the person coming up with it (based on considerations regarding the so-called "psychological creativity", which will be explained below). An idea is *surprising* if it

[1] Gaut, B. (2010).

[2] Boden, M. (2004); Boden, M. (1998).

[3] Boden, M. (2004), pp. 1, 11, 13.

[4] Boden, M. (2004), pp. 1, 35, 123, 269; Boden, M. (1998), p. 347.

[5] Boden, M. (2004), pp. 88-124, 269.

has not or could not have arisen so far under the constraints, i.e. from the application of the rules, defining the respective domain.[6] For an idea to be *of value*, it must be "*useful, illuminating, or challenging in some way*."[7]

In Boden's view, creativity is largely predicated on expert knowledge of and expertise in the domain being explored and transformed. Without extensive knowledge and skills in the respective area, the creator would not be able to systematically identify, map out, search and transcend the respective domain-relevant generative principles and conceptual structures. Boden elaborates on instances when the creator (e.g. Mozart, Darwin, and Shakespeare) knew his subject matter extremely well and was, as a result, able to build upon the state of the art in order to come up with new ideas. Greater insights allow for better understanding of the respective structured conceptual space ("the frame", as Boden further dubs it), and the resolved and unresolved questions ("filled/unfilled slots") therein. They enable seeing the solution in its entirety at once.[8] Expert knowledge also underpins the transformational potential of an idea.[9]

Boden explains human creativity by drawing comparisons and distinctions between generation of ideas in the human mind and computation in computers. Thought-processes are compared to problem-solving programmes. Exploration and human heuristics used therefor are likened to search by computers, and search-spaces/mapping of conceptual spaces – to search-trees. Data and action-rules in programmes are considered to form generative systems, which are similar to generative principles employed in the course of exploration and transformation of conceptual structures. Associative thinking is considered to resemble semantic nets in AI.[10] Connectionist systems (such as pattern-matching, pattern-completion, analogical pattern matching, and contextual memory) underpin both neurological processes and the operations of algorithms. In Boden's view, they are also essential to combinational creativity (see below for further details of this type of creativity).[11]

1.1.2 The Recombination Theory of Creativity

David Novitz provides an alternative theory to the computational theory of creativity.[12] He considers the latter underdescriptive due to its over- and under-inclusiveness at once.

On the one hand, the computational theory overstates the importance of generative principles and conceptual space as a benchmark, against which acts to be assessed and indeed asserted as creative. Since, as Novitz notes, inventions may come about without overcoming constraints and/or transforming conceptual spaces. As a result, the computational theory understates the creative propensity of acts that

[6] Boden, M. (2004), pp. 2-3, 40-53.

[7] Boden, M. (2004), p. 41.

[8] Boden, M. (2004), pp. 22, 35, 268-269.

[9] Boden, M. (2004), pp. 22, 123, 268.

[10] Boden, M. (2004), pp. 88-124, 131.

[11] Boden, M. (2004), pp. 125-133.

[12] Novitz, D. (1999).

do not transcend generative principles and transform conceptual spaces. Since there are such acts that build up — e.g. by way of recombination — on existing ideas, techniques, etc., and still result in novel, surprising and valuable outcomes, i.e. are creative.[13]

On the other hand, by overemphasising the importance of transformation, the computational theory — Novitz contends — casts a too wide net and qualifies as creative acts that may overcome existing generative principles but do so by mere trial and error and no true ingenuity involved.[14]

As a result, in Novitz's view, the computational theory may capture and qualify some acts as creative and others as not, irrespective of their creative parity.[15]

To remedy these deficiencies of the computational theory, Novitz advances a theory on the count of which an act is creative if it involves (i) "*intentional or chance recombination*" of ideas, techniques, etc. which recombination is "*subsequently deliberately used or employed*" in ways that (ii) bring about an outcome/output that "*is (or would have been) surprising — hence, not predicted by —- a given population*" and (iii) are "*intended to be, and are potentially, of real value to some people*."[16]

1.1.3 The Experiential Theory of Creativity

Bence Nanay attempts to reconcile the computational and recombination accounts of creativity by striking a common ground at a different — not functional but experiential — level of the creative processes. Nanay considers that what distinguishes creative from noncreative acts is not the manner of their implementation/manifestation (e.g. radical transformation, or recombination) but the experiences that the creator goes through.[17] An idea is creative if the creator takes it in her mind as something that has not been possible before and the idea in question has indeed not been possible before (and not merely an idea learned from somebody else).[18] I.e. creative processes are experienced "*in a certain distinctive way*", in which noncreative processes are not.[19]

Nanay's main points is that, although neural processes are "*causally responsible*" for creative ideas, it is the experiential level that the latter need to be analysed on (and not the functional one) in order to set creative from noncreative acts apart. In particular, acts are creative when accompanied by "*the experience of creativity*".[20] This is the common feature that unites different mental processes into a category of creative acts. Nanay does not argue for a particular set of necessary and sufficient conditions for creativity. He

[13] Novitz, D. (1999), pp. 71-74.

[14] Novitz, D. (1999), pp. 74-76.

[15] Novitz, D. (1999), pp. 75-76.

[16] Novitz, D. (1999), p. 77.

[17] Nanay, B. (2014), p. 18-21.

[18] Nanay, B. (2014), pp. 23-24.

[19] Nanay, B. (2014), pp. 18-21.
Nanay, B. (2014), p. 26.

[20] Nanay, B. (2014), p. 24, 30.

admits that the feature of creativity listed above may not be exhaustive. However, he claims that the concept of creativity is so vast and complex that it may very well not be possible to pin down all the conditions that define it.[21]

2.2 Types of creativity: within and beyond any perceivable boundaries

Boden distinguishes between:

- Based on the novelty of an idea to its creator and/or historically: psychological creativity ("P-creativity") and historical creativity ("H-creativity"). P-creativity presupposes generating a surprising and valuable idea that is new to its creator. H-creativity involves coming up with a surprising and valuable idea that arises "*for the first time in history*".[22]
- Based on the ways for generating ideas: combinational, exploratory and transformational creativity. Combinational creativity involves "*novel (improbable) combinations of familiar ideas*". Exploratory creativity presupposes coming up with novel, surprising valuable ideas by exploring structured conceptual spaces. Transformational creativity "*involves the transformation of some (one or more) dimensions of the space, so that new structures can be generated which could not have arisen before*".[23]

Other authors have drawn distinctions between objective and subjective creativity that broadly go into the direction of Boden's distinction between P-creativity and H-creativity.[24]

2.3 Originality, Randomness, Intention, Autonomy

The various attempts for devising a coherent theory of creativity have inevitably involved a more considerate thought being given to the correlation with originality and the role of intention, randomness, and autonomy in the creative process.

Creativity and originality are two different breeds of cattle. Creativity is considered to be a feature of mental process rather than of an entity (work, product). As creative is considered an act that is not mechanical. Originality is deemed to be a feature of "*objectively observed entities*" and not processes.[25] As original is regarded an output that is not derivative but unique, "*first of its kind*" historically.[26] An original idea may still not be creative while the characteristics of a mental process as creative does not necessarily testify of the output as original or not.[27] This distinction is important for the purposes of delineating further the boundaries and essential components of creativity, and for the discussion regarding rewarding creativity in Chapter 5.

---

[21] Nanay, B. (2014), p. 26.

[22] Boden, M. (2004), pp., 2, 43.

[23] Boden, M. (2004), pp. 3-6; Boden, M. (1998), pp. 348-349.

[24] See Nanay, B. (2014), p. 18 for an overview of those views.

[25] Nanay, B. (2014), p. 19.

[26] Gaut, B. (2010), p. 1039.

[27] Nanay, B. (2014), pp. 18-19; Gaut, B. (2010), pp. 1039-1040.

Randomness admittedly plays a certain role in the creative process. It also has a bearing on the qualification of acts as creative in a two-fold manner. On the one hand, not all creative processes involve pre-conceptualization and neat execution of an aesthetic vision or a scientific concept. Some degree of chance or serendipity may play out in the creative process and, yet, output still be ascribable to its creator and not mere occurrence of circumstances.[28] On the other hand, determinisms in processes may deprive the outcomes from the quality of being surprising, and, thus, dismiss their creativity.[29] However, outcomes that are entirely due to chance or serendipity are not to be considered as creative either, since creativity is considered to presuppose agency — be it human or other — as creative processes are actions and not mere byproduct of, e.g., luck. [30]

Creativity is also conditioned on the creator's ability to exert autonomy in the course of the creative process by evaluating aesthetic or scientific qualities of a work and changing, as need be, its features or generative standards applied.[31] Lack of autonomy is considered to result in what Boden calls "*automatism*" in the creative process.[32] In which case input predetermines output and preempts agency, novelty and surprising-ness.

Intention does not appear to be a *conditio sine qua non* to creativity. Intentional states are involved in the so-called "active creativity", i.e. when creators engage in deliberate creative pursuits. "Passive creativity" does not require intention as creative ideas emerge without any specific pre-conceptualizing or plan, merely "on the go".[33]

3 Can AI be creative?

It appears prudent to purport to answer this question by reference to the three theories discussed above in order to ascertain if AI can live up to the conceptual rigour of creativity that they introduce.

On the computational theory of creativity, it appears conceptually plausible that AI systems can be creative. The emergence and advancement of machine learning and deep learning in particular have rendered possible the exploration and transformation by AI of conceptual spaces in all its forms. In a paper of 2010, Boden reaffirms this (then) emerging tendency by explicating how various algorithm-based programmes meet the criteria for combinational, exploratory and even transformational creativity.[34] Generative Adversarial Networks (GANs), Convolutional Neural Networks (CNN), Neural Style Transfers (NST) and Artificial Intelligence Generative Adversarial Networks (AIGANs) have facilitated novel and imaginative art

[28] Boden, M. (2004), pp. 234-237; Gaut, B. (2010), pp. 1040.

[29] Boden, M. (2004), pp. 238-242.

[30] Gaut, B. (2010), p. 1041.

[31] Jennings, K. (2010), p. 489-491.

[32] Boden, M. (2004), p. 33.

[33] Nanay, B. (2014), pp. 30-31.

[34] Boden, M. (2009), pp. 25-31. Novitz and Nanay draw similar conclusions regarding feasibility of computer creativity on Boden's account. Novitz, D. (1999), p. 71. Nanay, B. (2014), p. 21.

generation.[35] AIGANs in particular have transcended pre-existing conceptual and style limitations stemming from the training sets, and equipped AI with capabilities to learn generative principles, apply them and deviate from them.[36]

Evolutionary algorithms, GANs and AIGANs appear to largely resolve also the concerns about lack of creative autonomy in AI while also instilling a fair amount of randomness in the system so that output may not be consider deterministic and, hence, a function of the underlying data/image sets and/or programmer's own creativity.[37] (It is another matter to what extent determinism as a theoretical concept is grounded given that quantum mechanics appears to suggest that indeterminacy might rather prevail.[38]) Computational models of creative appreciation, autonomous evaluation and change have been proposed by several authors,[39] generally admitted as possible in principle,[40] and tested in practice (e.g. in the project "The Painting Fool").[41]

On the recombination theory, creativity also appears within reach for AI. Recombination as output from the deployment of statistical and logical models is at the heart of programmes such as JAPE. The requirements that the recombination is "*deliberately used or employed*" in ways that are surprising and of value could be considered met by, for example, AlphaGo (when coming up with its infamous move No. 37 against Lee Sedol)[42] and by OpenAI's DALL·E (when creating images from text captions[43]).

The experiential account of creativity appears to require from the creator a certain level of awareness of the creative nature of the process, i.e. consciousness about the creative, in order for that process to qualify as such (creative). Which inevitably invites the question of whether AI systems are (at least hypothetically) capable of consciousness, i.e. of having mental states such as qualia, etc. Which, in turn, raises a number of philosophical and psychological questions regarding the essence of consciousness, mental state and mind in principle and the framing of those concepts. All questions that contemporary philosophy and science still do not have categorical answers to and which this essay is not meant to address in detail. Yet, it is worth setting out here several general considerations.

First, philosophy allows for possible equivalence between mental states and physical (brain) states (type-identity theory[44]) and for the multiple realisability of mental states by physical states (functionalism[45]).

[35] Cetinic, E., and She, J. (2021), pp. 6-10.

[36] Elgammal et al. (2017), pp. 5-6.

[37] Cetinic, E., and She, J. (2021), pp. 9; Elgammal et al. (2017), pp. 5-6.

[38] Rovelli, C. (2021), pp. 57-61.

[39] Jennings, K. (2010), p. 492-499.

[40] Cohen et al. (2012), pp. 97-105.

[41] Colton, S. (2012), pp. 7-25.

[42] Du Sautoy, M. (2019), pp. 18-44.

[43] Ramesh et al. (2021).

[44] Philosophy of Mind (2021a).

[45] Philosophy of Mind (2021b).

Hence, if neural networks emulate the human brain increasingly authentically, it can be reasonably argued that, by being identical to or a function of mental processes, physical processes at play in AI are representations of possible mental ones. I.e. that mental states in AI are not logically impossible.

If we are still concerned that, in order for AI to be creative, it should -- but does not -- exhibit mental states, philosophy may hold the answer to that concern as well. Epiphenomenalism advances plausible explanations as to why mental states, such as consciousness, may exist and still not have specific physical projections.[46] Moreover, a number of further arguments can be put forward as to why mental states may exist even without being causally efficacious.[47]

Those two sets of considerations come to the defence of the thesis that it is not logically impossible and, indeed, theoretically implausible that AI systems may be capable of consciousness even if it has no manifestations in the outside world and, hence, is not evident to us.

Third and most importantly, contemporary computer science considers machine consciousness within reach. Russell and Norvig deem that "[*i*]*ndividual aspects of consciousness – awareness, self-awareness, attention – can be programmed and can be part of an intelligent machine*."[48] Selman also expects that multi-agent systems will develop consciousness in order to be able to interact with each other.[49] Bostrom depicts a convincing picture of the possible emergence and rise of strong AI, with requisite attributes of superintelligence and consciousness.[50]

Were sceptics' views[51] regarding the untenability of machine consciousness nevertheless to be credited, hybrid forms of AI may still prove them wrong. Recent research into the possible interplay between computer and neuroscience suggests that neuron-silicon hybrid computing chips (the so-called "brains-on-a-chip") have more efficient computing capabilities, and are able to process information regarding the surrounding environment and act autonomously upon it.[52] Which appears to hold a promise for developing artificial general intelligence and, thus, artificial sentience.

4 Can AI (re)define creativity?

The considerations above uphold the possibility for AI creativity. It cannot be dismissed bluntly. Yet, our intuitions still speak strongly in favour of creativity as a domain reserved for humans only. How then these outcomes change (if at all) our predispositions to the notion of creativity? In my view, they prompt us to

[46] Jackson, F. (1982), pp. 283-289. Philosophy of Mind (2021c).

[47] Fessenko, D. (2021), pp. 1-4.

[48] Russell, S., and Norvig, P. (2021), p. 1037.

[49] Perry, L. (2021).

[50] Bostrom, N. (2016), pp. 26-61, 64-71110-120.

[51] Searle, J. (1980), pp. 6-8; Brooks, R. (2018); Wooldbridge, M. (2020), pp. 305-317, 327-334.

[52] Kagan et al. (2021).

unravel (soon!) the true essence and psychological traits of creativity and pinpoint the specific determinants that qualify a process as "creative".

Boden admits that creativity "*involves not only a cognitive dimension (the generation of new ideas) but also motivation and emotion*".[53] Novitz makes a convincing case as to why mere mechanical exploration/recombination/transformation of ideas on a "trial-and-error" basis does not constitute a creative act. Nanay appears to have a point by appealing to the experiential aspects of the mental process, as, indeed, any accidental effort or relentless trial-and-error exercise may otherwise easily cross the mark for novelty, surprising-ness and value if taken at par. Hence, it seems justified to allow for some further, mental components as essential to creativity. Otherwise, would the creative process have been purely computational/functional, then, logically, creativity is completely attainable for entities with control centres emulated on human brains (such as modern AI). Which trumps our intuitions above.

Awareness of the creative nature of an ongoing process and experiencing it as such (i.e. having the raw feel of "creating") appear to be two such further essential mental components. They play out in an intellectual endeavour in order for it to stand out as "creative" rather than mere occurrence of circumstances, or mechanical effort. Consciousness and experiential states appear also logically necessary for creativity for two further reasons. Consciousness and raw feel of the creative process constitute the real-time connection to the latter and ensure perception of its course and ability to act on it. I.e. they facilitate the possibility for action, i.e. for agency, in the creative process. Second, consciousness and raw feel enable also aesthetic appreciation, evaluation, and deviations, and, in this way, autonomy. Otherwise, the ability for adjustments would be preempted and reactiveness would set in.

The key qualifying factor for creativity appears, however, to be the intent to instigate creatively (i.e. through a creative process) a change to the status quo. Such a propositional attitude gives a direction to the creative endeavour. Intent to instigate change creatively practically materialises and manifests the agency and autonomy involved. It channels them in such a way so as to premise novelty and value. Absent such intent, mechanical recombination or trial-and-error exercises (such as in Goodyear's case) may plausibly be considered creative (although not all pundits agree with that[54]). By the same token, the mere deployment of the brute computing force of artificial general intelligence (assuming it exhibits consciousness and experiential states, as per above) may still rank as creative.

5 Societal benefits and rewards

If the prospect of AI creativity convinces us of the need to fine-tune the overall concept of creativity along the lines above, what wider societal implications this may have?

[53] Boden, M. (1998), p. 347.

[54] Novitz, D. (1999), pp. 75-76 ; Nanay, B. (2014), p. 25.

5.1 The appropriate conceptual and legal framework for rewarding creativity

The first and most obvious (at least to lawyers) implication is the need to revisit, conceptually and legally, also the approach to ascribing and rewarding creativity.

The existing legal framework rewards creative endeavours by reference to the *originality* or *novelty* of the end product specifically. The Berne Convention for the Protection of Literary and Artistic Works[55] grants protection and rights to reward with respect to literary or artistic works that are "*original*".[56] The World Trade Organisation's Agreement on the Trade-Related Aspects of Intellectual Property Rights[57] warrants patent protection for inventions "*provided that they are new, involve an inventive step and are capable of industrial application*".[58] Hence, both legal regimes uphold the qualitative features of the *product* as qualifying factors for protection and reward. None of these legal standards accentuates on agency or the capacity of the inventor (e.g. as a natural person) as a qualifying factor for granting legal protection and rights to commercialization and associated reward. Both legal frameworks also set the standards for national protection in the member states that are signatories to those international treaties. As a result, those legal standards are widely adopted among developed and developing countries.

However, the bar of novelty/originality of the output is not insurmountable for AI. As discussed above, AI systems can create works that are novel and/or original. This creates a host of questions for deliberation among scholars, artists/inventors, industry, policy- and lawmakers. Among them are: are standards by reference to the originality/novelty of the work/invention still adequate? Should they not be upgraded with regard to the role of the creator in the process? If not, are we ready to recognise a legal status of "creator" to AI? And in what set-up – as a co-creators always (alongside the programmer) or possibly in its own right? If such status is recognised, who reaps the rewards for AI's creations – the programmers of the AI systems, the corporations that own them, or the general public?

The latter question is particularly acute given that algorithms are trained and tested on vast sets of data, some of which include copyrighted materials (e.g. literary works, images of artworks), personal data (e.g. regarding individual aesthetic preferences) or inferences from such. If AI creative capabilities are (em)powered by copyrighted materials or personal data, further questions regarding the fair distribution of benefits from a creation arise. For example, should – and if yes, how – the general public benefit from inventions / creative works of AI systems if troves of personal data are the key component that fuelled and informed creative choices? If yes, on what basis – e.g. by way of open source access, open access licences,

[55] Berne Convention for the Protection of Literary and Artistic Works, of September 9, 1886, as last revised in Paris on July 24, 1971 (the "**Berne Convention**").

[56] Per argument from Art. 2(3) and (8), and Art. 14bis et seq. of the Berne Convention.

[57] Agreement on the Trade-Related Aspects of Intellectual Property Rights (the so-called "**TRIPS Agreement**"). The TRIPS Agreement is Annex 1C of the Marrakesh Agreement Establishing the World Trade Organisation, signed in Marrakesh, Morocco on 15 April 1994.

[58] Art. 25 of the TRIPS Agreement.

or other forms of availability to the wider public, or sharing the monetary benefits in a more structured way which allows for reinvestment and pursuit of public interest causes?

Some of these issues have been debated — and the respective legal standards tested in litigation — before US, UK and European Union authorities and courts.[59] Yet, the jury is still out on the ultimate resolutions. Agency and autonomy are morally and legally relevant for assigning responsibility.[60] They should be morally and legally relevant also for assigning rights. In the case of creativity, consciousness, raw feels for creating and intent to creatively constitute change arguably underpin agency and autonomy. Therefore, legal standards should find a way — through law or precedent — to accommodate them as the morally and legally significant elements for ascribing creativity and associated reward.

6 Conclusions

AI can and is already redefining the concept of creativity. Or at least our perception of this psychological phenomenon. This happens in a rather typical way for situations involving intelligent machines — by us (humans) realising that algorithms can do what we can, and that recognition and associated benefits therefor might naturally be in store for them as well. The time appears ripe for the respective conceptual and legal frameworks to evolve and account for consciousness, experiential states and propositional attitudes — being the underlying causes of agency and autonomy — as qualifying factors for bestowing creativity. As much as mental states, agency and autonomy are the triggers for moral and legal responsibility, they should be among the key determinants for assigning legal rights and ensuing rewards as well. As advanced AI emerges, it may be meeting those criteria and formally qualify for the status of an "inventor"/"author". Then, we would need to resolve also the implications of granting AI such a status and of the fair access to and distribution of the benefits from AI's creations.

10 January 2022

*

* *

[59] Most notably, in a recent series of patent applications by Dr. Stephan Thaler designating the AI system DABUS as inventor. Resolution by the European Patent Office available at: https://www.epo.org/law-practice/case-law-appeals/communications/2021/20211221.html. Decision by the England and Wales Court of Appeal available here: https://www.bailii.org/ew/cases/EWCA/Civ/2021/1374.html. Decisions in similar applications by Dt. Thaler rendered by courts also in the US, Australia and South Africa.

[60] Talbot, M. (2011), slides 4-31.

References:

1. Boden, M. (1998). *The Creativity and Artificial Intelligence*. Artificial Intelligence 103 (1998), pp. 347-356;
2. Boden, M. (2004). *Creative Minds: Myths and Mechanisms*, 2nd ed. Oxon, Routledge;
3. Boden, M. (2009). *Computer Models of Creativity*. AI Magazine FALL 2009, pp. 23-33. Last accessed on 4 January 2022 at: https://doi.org/10.1609/aimag.v30i3.2254 ;
4. Bostrom, N. (2016). *Superintelligence: Paths, Dangers, Strategies*. Reprint edition. Oxford, Oxford University Press.
5. Brooks, R. (2018). *Strachey Lecture – Steps towards Super Intelligence*. Last accessed on 6 January 2022 at: https://podcasts.ox.ac.uk/strachey-lecture-steps-towards-super-intelligence;
6. Cetinic, E., and She, J. (2021). *Understanding and Creating Art with AI: Review and Outlook*. arXiv preprint arXiv:2102.09109v1;
7. Chalmers, D., ed. *Philosophy of Mind: Classical and Contemporary Readings*, 2nd ed. Oxford, Oxford University Press;
8. Cohen, H., Brown, D., Brown, P., Galanter, P., McCormack, J., d'Inverno, M. (2012). Evaluation of Creative Aesthetics. In McCormack, J., d'Inverno, M., ed., *Computers and Creativity*, 1st ed., pp. 95-111.
9. Colton, S. (2012). The Painting Fool: Stories from Building an Automated Painter. In McCormack, J., d'Inverno, M., ed., *Computers and Creativity*, 1st ed., pp.3-36;
10. Elgammal, A., Liu, B., Elhoseiny, M., and Mazzone, M. (2017). *Can: Creative adversarial networks, generating" art" by learning about styles and deviating from style norms*. arXiv preprint arXiv:1706.07068;
11. Fessenko, D. (2021). *If Mental States are Causally Inert Do We Have Any Reason to Believe They Exist?*. Forthcoming.
12. Gaut, B. (2010). *The Philosophy of Creativity*. Philosophy Compass 5/12: pp. 1034–1046. Last accessed on 23 December 2021 at: https://www.sfu.ca/~kathleea/docs/The%20Philosophy%20of%20Creativity%20-%20Gaut.pdf
13. Jackson, F. (1982). Epiphenomenal Qualia. In Chalmers, D., ed., *Philosophy of Mind: Classical and Contemporary Readings*, 2nd ed., pp. 283-289;
14. Jennings, K. (2010). *Developing Creativity: Artificial Barriers in Artificial Intelligence*. Minds & Machines (2010) 20:489–501. Last accessed on 23 December 2021 at: https://link.springer.com/article/10.1007/s11023-010-9206-y;
15. Kagan, B., Kitchen, A., Tran, N., Parker, B., Bhat, A., Rollo, B., Razi, A., Friston, K. (2021). *In vitro neurons learn and exhibit sentience when embodied in a simulated game-world*. BioRxiv Preprint. Last accessed on 6 January 2022 at: https://doi.org/10.1101/2021.12.02.471005 ;
16. McCormack, J., d'Inverno, M., ed. (2012). *Computers and Creativity*. 1st ed. Berlin, Heidelberg: Springer Berlin Heidelberg.

17. Nanay, B. (2014). An Experiential Account of Creativity. In Paul, E., and Kaufman, S., *The Philosophy of Creativity*, pp. 18-35;
18. Novitz, D. (1999). *Creativity and constraint*. Australasian Journal of Philosophy, 77:1, 67-82. Last accessed on 4 October 2021 at: https://doi.org/10.1080/00048409912348811 ;
19. Paul, E., and Kaufman, S. (2014). *The Philosophy of Creativity.* Oxford, Oxford University Press;
20. Perry, L. (2021). *Bart Selman on the Promises and Perils of Artificial Intelligence*. Future of Life Institute. Last accessed on 5 January 2022 at: https://futureoflife.org/2021/05/20/bart-selman-on-the-promises-and-perils-of-artificial-intelligence/.
21. Philosophy of Mind (2021a). *Unit 3: Type-identity theory*. Short Online Course. The Department for Continuing Education, University of Oxford. Last accessed on 3 November 2021 at: https://michaelmas2021.conted.ox.ac.uk/mod/book/view.php?id=1530&chapterid=2363;
22. Philosophy of Mind (2021b). *Unit 4: Functionalism*. Short Online Course. The Department for Continuing Education, University of Oxford. Last accessed on 3 November 2021 at: https://michaelmas2021.conted.ox.ac.uk/mod/book/view.php?id=1530&chapterid=2363;
23. Philosophy of Mind (2021c). *Unit 8: Epiphenomenalism*. Short Online Course. The Department for Continuing Education, University of Oxford. Last accessed on 3 November 2021 at: https://michaelmas2021.conted.ox.ac.uk/mod/book/view.php?id=1530&chapterid=2363;
24. Ramesh, A., Pavlov, M., Goh, G., Gray, S. (2021). *DALL·E: Creating Images from Text*. Last accessed on 7 January 2022 at: https://openai.com/blog/dall-e/ ;
25. Rovelli, C. (2021). *Helgoland*. Allan Lane.
26. Russell, S., and Norvig, P. (2021). Artificial Intelligence: A Modern Approach. 4th global ed. Pearson;
27. Du Sautoy, M. (2019). *Creativity Code: How Ai is Learning to Write, Paint and Think*. Fourth Estate Ltd.;
28. Searle, J. (1980). Minds, brains, and programs. The Behavioural and Brain Science (1980) 3, pp. 417-457;
29. Talbot, M. (2011). *A Romp Through Ethics for Complete Beginners, Session Two: Freedom, knowledge and society: the preconditions of ethical reasoning*. Oxford Online Course Portal. Last accessed on 3 June 2020 at: https://trinity2020.conted.ox.ac.uk/mod/book/view.php?id=479];
30. Wooldbridge, M. (2020). *The Road to Conscious Machines. The Story of AI*. Pelican Books.