RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

Is Delulu the New Trululu?[1] Artificial Intelligence Hallucinations as Input in the Creative Process

Is Delulu the New Trululu? *Alucinações da Inteligência Artificial como Contributo para o Processo Criativo*

JOÃO G. PATRÍCIO

ICNOVA — NOVA Institute of Communication, Portugal
NOVA University Lisbon, School of Social Sciences and Humanities
a41371@campus.fcsh.unl.pt

Abstract

From Gods to mere mortals, creativity has been studied throughout the last decades as an intrinsic human capability of bringing something into existence. When computers are added to the equation, discussion arises. Numerous authors have defended the inexistence of computational creativity. However, if we consider the rise of Artificial Intelligence (AI), and focus on the cases of hallucinations in artistic creation, can AI hallucinations constitute a creative input into the creative process alongside a human agent? The present work proposes a literature review to understand what creativity means, if the phenomena of hallucinations can be considered creative by themselves and if they can serve the creative process. Here, it is shown that AI hallucinations can be creative and can be an input in the process of artistic creation under the artist's appreciation. This paper offers another point of view in favour of computational creativity, that aims to contribute to a fruitful interaction between artists and AI.

Keywords

creativity | creativity studies | computational creativity | artificial intelligence | artificial intelligence hallucinations

Resumo

Dos deuses aos meros mortais, a criatividade tem sido estudada ao longo das últimas décadas como uma capacidade humana intrínseca de trazer algo à existência. Quando se juntam os computadores à equação, surge a discussão. Vários autores têm defendido a inexistência de criatividade computacional. No entanto, se considerarmos a recente ascensão da Inteligência Artificial (IA) e nos centrarmos nos casos de alucinações na criação artística, poderão as alucinações da IA constituir

um contributo criativo para o processo criativo, a par de um agente humano? O presente trabalho propõe uma revisão da literatura para compreender o que significa criatividade, se os fenómenos de alucinações podem ser considerados criativos por si mesmos e se podem servir o no processo de criação. Neste artigo, mostra-se que as alucinações da IA são, à partida, criativas, e um contributo para o ato criativo. Este artigo oferece outro ponto de vista a favor da criatividade computacional, que visa contribuir para uma interação frutuosa entre artistas e IA.

Palavras-chave
criatividade | estudos de criatividade | criatividade computacional | inteligência artificial | alucinações de inteligência artificial

1. A History of highs and lows

Unravelling the Artificial Intelligence (AI) world means going back a few decades to the 1950s. The first discussions on the ability of thought in machines and the conceptualisation of AI started with Alan Turing with the straightforward question "Can machines think?" (1950, 433). Without denying the possibility of thinking machines, Turing proposes a forecast that involves machines passing his computational intelligence test with more than a 70% chance within 50 years. (*ibid*, 340). He also presents the concept of "Learning Machines", i.e. machines that can be trained to learn and create memories. Like the human intelligence development process from birth to adulthood, machines should go through "an appropriate course of education" similar to what children do (*ibid*, 352).

The following year, Turing stated that machines can simulate human behaviour. They would make mistakes and learn from past experiences, in a process he called "education" (1996, 257) which we nowadays call "training". Such machinery would have a *memory* (*ibid*, 258) and a *"random element"*, acting without full determination of its experiences (*ibid*, 259).

In 1956, the Dartmouth Summer Project outlined the Artificial Intelligence movement, inspired not only by Turing's "logical computing machines" (Turing 1992) but also by John von Newmann's modern central processing unit (CPU), Norbert Wiener's Cybernetics and Claude Shannon's Information Theory (Jansen 2021, 11-12). This event marked the beginning of the study of AI, which was characterised by distinct phases.

1 The terms "delulu" and "trululu" are etymological derivations of the words "delusional" and "true", and made part of a recent TikTok trend called "May all your delulu come trululu" https://www.nytimes.com/2023/11/23/style/delulu-tiktok-gen-z-millenials.html, consulted in February 26, 2024). The tittle reappropriated the terms by expressing a validation (trululu) of Artificial Intelligence hallucinations (delulu).

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

This first phase of research, during the 1950s and the 1960s, was a time of "overconfidence" in AI, not only among researchers but also among the media and the general public (Mitchell 2021, 2). Multiple experts made their predictions on the capabilities of AI. Since then, AI has gone through Winters and Springs of optimism until today. The 70s saw the first AI winter, with negative prospects and consequential cuts in funding. By the beginning of the 80s, a new hope arrived in the field, as Japan joined the AI-craze with its "Fifth Generation" project and the United States promoted the "Strategic Computing Initiative", both aiming to develop AI capabilities on computer systems (*ibid*). Such projects, among others at the time, revealed a few problems regarding the generalisation of conduct, meaning the developed systems worked well in specific contexts but failed to adapt to other situations. This led to a new AI-Winter by the early 90s.

This period is also mentioned as the "top-down" or "strong" approach in the Artificial Intelligence Movement (AIM) (Jansen 2021, 12). These were the times when researchers were moved by the idea of computers and machines that could only do what they were told to do, aiming to produce computers that were able to "think" and to "learn" (*ibid*). Computer programs are meant to simulate the human mind, empowering and even surpassing it. Such a time frame extended until the mid-80s, more or less synchronous to the end of the second AI-winter in research.

The 1990s witnessed a blossoming of interest in AI research. The leading cause was the rise of Machine Learning ML and predictive models (Mitchell 2021, 2). These advancements took inspiration from statistics and provided a solution to specific tasks, struggling to aim for generality.

Simultaneously, the AIM enters its second phase, the "weak" or "bottoms-up" approach. In this period, the development of expert systems, robotics, and commercial applications of AI is still ongoing (Jansen 2021, 12-13). However, the approach focuses more on biology and the brain's physical structure in this new phase. The goal is no longer to mimic a fully developed human brain, but instead to take inspiration from child development, namely on the actions of observing and learning (*ibid*,13).

By this time, another discipline caught the attention of the research community: Artificial Life (AL). This new discipline seeks to contribute to the evolution of conventional biology with the study of "man-made systems that exhibit behaviours characteristic of natural living systems" (Hayles, 1999, 232). It analyses other living organisms (not strictly humans) and tries to synthesise them into computers and artificial media. AL strives to design a new way of evolution of life on earth, where natural and artificial life will be difficult to distinguish (*ibid* 1999, 235). This may be responsible for some of the anti-computational creativity positions that will be addressed later on.

The optimism felt in the 90's research environment got even bigger in the 2010s when technological advances such as Machine Learning or Deep Neural Networks brought new and transforming functionalities into our daily devices (Mitchell 2021, 2-3). Most recently, tools in the domain of Generative Artificial Intelligence (Gen AI) like Open AI's ChatGPT, DALL.E, and MidJourney became part of our daily routines.

However, it is important to keep in mind that "the AI cycle of springs and winters is likely to continue" (*ibid* 2021, 3) due to several fallacies related to how AI is understood, assessed, and communicated to the public (4-7).

2. AI and Creativity

Understanding creativity in this context involves addressing the concept from the human and computational dimensions. This section will provide relevant literature regarding not only each of these two domains but also their association.

2.1. Human Creativity

Runco and Jaeger propose a standard definition of creativity that is based on two pillars: one stands for originality, the other for effectiveness (2012, 92): 1) originality in the sense of novelty, distinctiveness, and uniqueness; 2) effectiveness in the sense of usefulness, appropriateness, and economic value.

However, the notion of creativity has a long genealogy. In fact, creativity as such is a "modern concept and a modern value" (Mason 2003; Weiner 2000; Reckwitz 2017 *apud* Glăveanu and Kaufman 2019, 10). If we regard the etymology of the word creativity, we conclude that it comes from the Latin, and means to bring something into existence (Glăveanu and Kaufman, 2019, 14). Such attribute in Western History was reserved for deities. Humans could only produce *art*. When bringing something into actual existence, such practice was a mere reproduction of the divine will (*ibid*).

Only with the Renaissance and its Anthropocentrism did the concept of creativity pass from God to humans, a current of thought that was later reinforced during the Enlightenment due to "a new belief in the power of human reason and capacity to change the world [which] offered the foundation for a much more individual notion of creativity" (*ibid*, 19).

The limits of creativity are also changing due to its "democratisation" (*ibid*, 20). Once restricted to a certain elite, creativity expands to a wider population of artists and artisans as it shifts from something acquired through education to a personal trait (*ibid,* 16).

Due to technological advances, Weiner states that creativity can come "from anyone, anywhere, at any time", merging "information, entertainment, technology, and art", thus bringing new ways of communicating and being (2000, 107).

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

2.2. Computational and AI (-assisted?) Creativity

Ada Lovelace was the first person to question the possibility of creative computers. She concludes that a computer can't be creative as it doesn't have a pretension to be so; it only performs what it is asked to do (Turing 1950).

Jefferson highlights the speed of electronic computing machines and how their components render the machine "a thousand times faster" than our brains (1949, 1107). However, he disclaims that whenever one finds some resemblance in computers with the human nervous system, such resemblance is not entirely coincidental with the human system (*ibid*, 1108). He furthermore marks a division between humans and machines:

> "Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain — that is, not only write it but know that it had written it." (ibid 1949, 1110).

In the artistic domain, Galanter introduces the concept of generative art, which "cedes control to a system that operates with a degree of relative autonomy and contributes to or results in a completed work of art" (2009, 2). Such practices involve "randomisation in composition", "the use of genetic systems", are "constantly changing over time", and their works are "created by running code on a computer" (*ibid*). Concerning computational creativity, the author admits there are still a lot of unsolved mysteries. As an artist himself, Philip Galanter acknowledges the difficulty of grasping computational creativity is related to our incomplete understanding of consciousness and "the related phenomena of self-awareness and the experience" (*ibid*). Also, creativity is a quality of complex adaptive systems that, to keep their integrity, adapt their nature to their context through actions of creativity (*ibid*, 16). Humans are an example of that. Their existential integrity is kept by maintaining their "congruity" with their social expectations. Their actions of creativity are always influenced by a social context. Culture can be an example of a social context (*ibid*). Such adaptative behaviour can be present in lower life forms and simultaneously be intelligent and creative. As long as a computer can be regarded as a complex adaptive system, "it can also be considered creative" (*ibid*).

2.3. AI Creativity

When asked, "can AI be creative?" Joler and Pasquinelli affirm that it can't since the process of Machine Learning (ML) is only based on training data, meaning that the detection of styles is constrained to what the input data contains, thus, rendering impossible what the author terms the "prediction" and "generation of the new" (2021).

On a similar line, Runco states that "AI can only produce Artificial Creativity". Narratives on AI that mainly focus on the outputs rather than the processes made people

believe that AI could be creative (2023, 1). He proposes either the extension of the current notion of creativity or establishing the distinction between AI's pseudo-creativity and the "authentic creativity of humans". Additionally, AI lacks a few conditions that are necessary for it to be creative. Despite AI's original outputs, originality or novelty doesn't always mean creativity, as it lacks intrinsic motivation and mindfulness. Authenticity is another aspect lacking in AI: a feeling of self-acceptance and self-expression regardless of the other's expectations. In AI "there is no self to express, so no possibility of authenticity" (3). AI's output comes from the process of scrapping information from the web, hence the term 'artificial creativity'" (6).

The same term was used by Carvalhais and Lee, who defined it as a sub-sector of AI that aims to understand and replicate human processes of inventiveness to develop creative computational systems (2022, 72). Such a domain comes as a solution between antagonistic views on whether creativity is strictly human or whether there can be some computational equivalent. The authors also pinpoint that art and creativity are not interchangeable, and in a context where machines are presenting more "agency, autonomy, and creativity", computational creativity can be "accessed" through computational art. Computational art raises another question, for these artistic practices encompass cases where computers can be mere tools or are the primary medium (73).

Broeckmann reflects on what an artist is and if computers can be included. The meaning of *being an artist* has been contested since the beginning of the 20th century, and "machine artworks" won't be responsible for the death of art, but will instead function as a vehicle of "continuous transformation of sense-making that we tend to categorise as 'art'" (2019, 3). Moreover, the Myth of the Machine, which places machines as a demonic entity that will eventually gain more autonomy, and consequently control humans, needs to be unmasked (4-5)—such mythological narrative results in fear and self-protection against machines.

2.4. A critical view on computational creativity

Margaret Boden introduces a different perspective on computational creativity compared to most of the aforementioned authors. To answer Lovelace's question on computational creativity, Boden (2004) presents four arguments that support an alternative perspective. For each argument against computational creativity, she proposes a possible answer.

2.4.1. The "human-stuff" in AI

The brain-stuff argument presents the idea that materials such as neuroproteins can support intelligence and materials like metal or silicon cannot. To this Boden states that

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

such a position is "inconclusive": as we don't fully understand how neuroproteins support intelligence (and consequently creativity), we can't fully deny that non-biological materials can't operate in a way that enables creativity (*ibid*, 287-289).

Besides, when we think about some of the architectures behind today's computational creativity, we can make some parallelisms with the human brain. Neural Networks are an example: "A neural network is a neurobiologically-inspired computing system" (Del Campo 2020, 95). To regard these human-like systems as somewhat inferior is one of the common mistakes about AI highlighted by Leonel Moura. According to Moura (2023), people tend to regard AI as a mere tool, with the assumption that humans are a superior species, which tends to diminish AI's capabilities.

Additionally, once computers and AI aren't made of the same organic materials as humans, they can't be credited for their artworks because behind them "hides a human operator" who made the creative process possible (Joler and Pasquinelli 2020). In this case, we must address the question of hallucinations. Hallucinations can be unexpected. Jefferson points out that computers demonstrate "spontaneous functional faults", producing an endless loop instead of proceeding with the programmer's intentions (1949, 1109).

In this case, if hallucinations are an example of deviation from the primary intention of a programmer, the question arises as to whether the authorial credit for the output of the hallucination should be *partially* given to the machine or not.

2.4.2. The intentionality issue

Let us return to Margaret Boden's answers on computational creativity. She introduces the empty-program argument, which states that "all the symbols dealt with by a computer program are utterly meaningless to the computer itself" (2004, 287). She uses John Searle's Chinese room experiment, which supports this argument. In Searl's vision, computers only process syntax, but no semantics, and all they do is process information without comprehending its meaning (1980, 354). Boden replies by stating that computer programs act with proper hardware, which allows the machine to enter into contact with the world and provides some form of semantics, i.e. the computer is not only processing symbols, it is interacting with the world and reacting to it:

A programmed instruction, then, is not merely a formal rule. Its essential function (given the relevant hardware) is to make something happen. (Boden 2004, 292).

Furthermore, in Searle's experiment, the subject may not understand a word of Chinese, but he understands English, in which the rule book was written to combine Chinese characters. In this sense, there are indeed, multiple operations that engage the realm of semantics to make the machine work properly (293).

Regarding the Lovelace Objection (Turing 1950), if hallucinations are an unexpected output, then computers do not always strictly execute what is demanded from them. Furthermore, we can refer to Margaret Boden's typology on computational creativity

(2004, 3-6), namely transformational creativity, which involves going beyond a particular bounded search space. Boden considers the possibility of computers going beyond the initial expectations of the programmers by transgressing their possibility space.

2.4.3. A matter of (AI) consciousness

The consciousness argument states that no computer displays sufficient consciousness to be creative (Boden, 2004, 287). Nonetheless, the process that leads to novelty is indeed unconscious, as has been proved by "reports given by artists, scientists, and mathematicians" (294). Creativity needs a type of consciousness, namely self-reflective thinking, which several computer programs (like AM by Douglas Lenat, DALTON, ARCS-ACME, and COPYCAT) provide in their own way (295). However, this self-reflective evaluation that is key for creativity is just one of many possible definitions of consciousness. As this term is still evolving, one can't fully understand what consciousness is. In these terms, Boden states that we can't be sure that no computer could ever be conscious (296; Galanter 2009, 15).

2.4.4. The non-human argument

The non-human argument states that an intelligent computer is morally absurd (Boden, 2004, 287). According to Boden, this argument reveals an idea of *impertinence* of machines: "we've admitted in the past human artists challenging our perceptions and aesthetic conventions, but we find it difficult to tolerate such impertinence from a computer program" (297).

Holding on to this position would mean that, regardless of technological advances, we would also have to stick to moral principles and continue disregarding computational creativity. However, if these new computers became more visually appealing, like "encased in fur" with big eyes, we could become more morally acceptant of computational creativity (299).

3. Beyond the Expected: AI Hallucinations

Joler and Pasquinelli (2021) refer to the importance of "critical inquiry studies" of how technology "breaks", in particular with the study of hacking. The present inquiry will focus on a different situation, namely on what can be considered a technical failure.

When using a computational system, it can "express itself" through bugs, glitches, or other inaccurate outputs, which disturb the intended stability of information (Carvalhais and Lee 2022, 76).

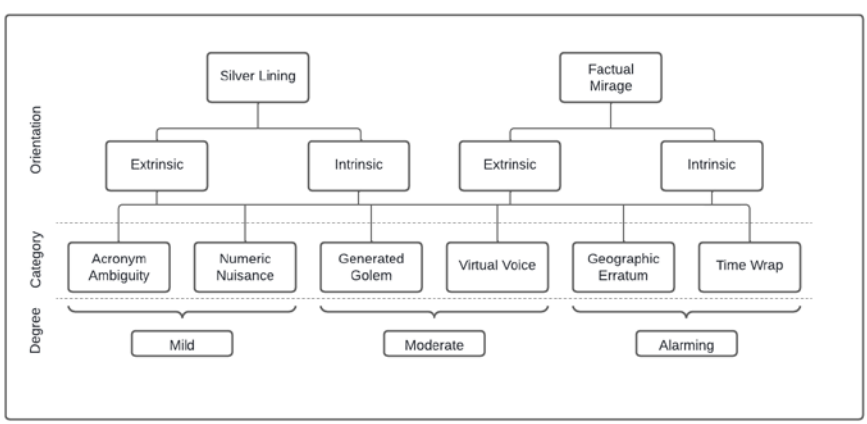
RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

In this context, the notion of "hallucination" must be introduced. In the realm of AI, a hallucination is generated content that does not correspond to real facts, which can result in inaccurate outputs (Rawte et al. 2023, 2542). There are two types of hallucinations (2543-2546):

— Factual Mirage: Occurs when given an accurate and factual prompt to a Large Learning Model (LLM). This can be divided into:
— Intrinsic Factual Mirage: The LLM generates a correct answer with supplementary information.
— Extrinsic Factual Mirage: The LLM diverges from factuality.
— Silver Lining: when given a factually incorrect prompt, the LLM engages in that falsity, generating an elaborate narrative about it. It can be classified as:
— Intrinsic Silver Lining: the generated narrative isn't convincing.
— Extrinsic Silver Lining: the generated narrative is convincing and persuasive.

These can be further categorised as:

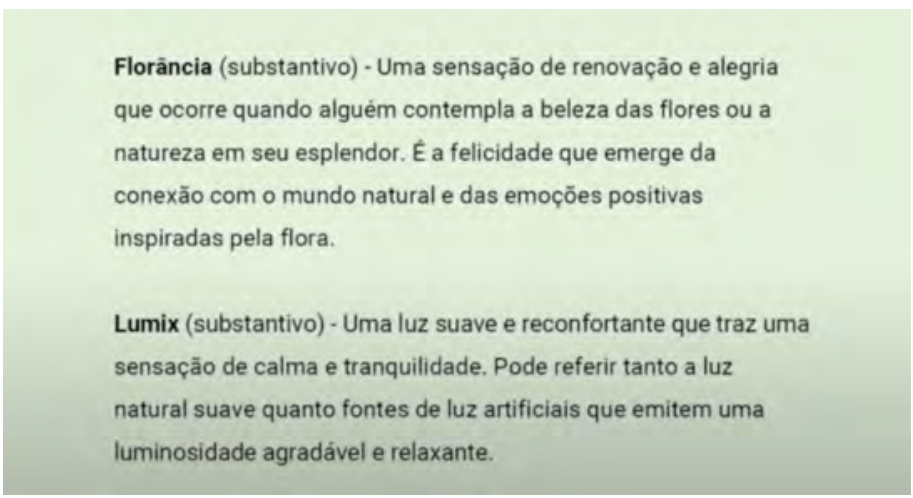
— Numeric Nuisance: The LLM provides incorrect numbers.
— Acronym Ambiguity: The LLM gives an incorrect expansion for an acronym.
— Generated Golem: The LLM generates a fake person regarding a past event.
— Virtual Voice: The LLM generated false statements from real or fake personalities.
— Geographic Erratum: The LLM generates a false location of an event.
— Time Wrap: The LLM mixes up different events from different times.



—
Figure 1
Hallucination typology adapted from Rawte et al. (2023, 1) | © Researchgate

The European Ambassador for Creativity and Innovation of the European Commission, Leonel Moura (2023), gave us an example of a deliberate use of AI hallucinations. The Portuguese artist asked ChatGPT to make up two words in Portuguese that would make sense. The AI tool retrieved two words in an occurrence that, applying the above typology, would be described as an Extrinsic Silver Lining.

Although the above list of hallucinations is tied to situations of AI-generated text, the same situation can occur with AI-generated images. Del Campo et al. provide a case study from Architecture. The project consists of a Robot Garden, a testing ground for robots (2020, 94). In terms of methodology, the researchers used a Neural Network, a computing system composed of layers of processing intersections called neurons, where information flows and generates representations of attributes (ex. columns, arches, fountains...) (*ibid*, 95-96). When the Neural Network was fed with satellite images of the project's construction site and it was asked to transfer the given 2D format into a 3D model, it gave back a "novel view" of the architectural features present in training data. It was described as a "hallucinogenic dream of a machine trying to see these features in the landscape" (105). It was also classified as a "successful attempt to use machine hallucinations, based on architectural imagery" (106).



—
Figure 2
A print screen from Leonel Moura's presentation with the two ChatGPT-generated terms and corresponding meanings. (Moura 2023) | © APDSI

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

Marks refers to the Lovelace test and recognises the possibility of a machine expressing creativity if its performance goes "beyond the intent of the programmer" (2022 *apud* Runco 2023, 3). If the programmer can't explain such "ostensible creativity", then it is safe to admit that the machine produced an idea (4).

When asked about how to deal with AI's hallucinations, Leonel Moura highlighted that, as an artist, these were the best part of ChatGPT (Moura 2023): "it is when the machine's creativity is most evident"[2] (2023). He additionally states that the further the machine diverts from truth and reality, the better because at that moment it will be inventing. However, he advises that this cannot be applied, for instance, to the realm of scientific inquiry. Moura likewise says that just as we can't deterministically trust a machine, we wouldn't deterministically trust the best professor, for they can also commit mistakes.

By addressing Leonel Moura's words with existing literature, we ask: *Can AI hallucinations provide creative input to the* creative *process alongside a human agent?*[3]

After the current section of the literature review, five critiques of computational creativity were chosen to be discussed under the eyes of various authors. Each point of discussion will address as much of the theme of AI hallucinations as possible to find hints that may provide answers to our initial inquiry questions. Finally, I will be sharing my conclusions on the subject.

4. Discussion

> "Glitches aren't flaws but gateways to creativity"[4]

This catchphrase is from a post on Glitch Magazine's Instagram account. It catches attention due to the power transfer that is given to malfunction. Going back to the discussion of Leonel Moura's words on AI hallucinations as evidence of computational creativity: are glitches and hallucinations proof of machines' artistic expression? At first glance, most of the literature gives us a 'no' for an answer.

2 Loose Translation.

3 The word "creative" is intentionally repeated in our research question. More than understanding whether hallucinations can be useful in the creative process, I want to discuss if hallucinations can be considered by themselves a creative expression of machines.

4 From https://www.instagram.com/p/C1pcGSjy2Kw/ (consulted on January 8, 2024).

4.1. The boundaries of computational creativity

Another critique against computational creativity refers to the fact that AI systems are bound by the datasets that were fed to them during the training process (Crawford and Paglen 2021; Joler and Pasquinelli 2021).

In a round of Q&A in the already mentioned conference, Leonel Moura says the following:

> All human art is derivative (...) No artist is capable of doing something completely out of any reference. We are all derivative. We have read, we have gone to school, everything's derivative. (Moura 2023)

We can thus expand the answer to this objection by introducing IBM's notion of AI hallucination[5], understood as the generation of inaccurate or nonsensical outputs. At a point in their explanation of these phenomena, it is mentioned that AI algorithms generate results that deviate from the training data or lack any recognisable patterns. As a result, computers go beyond the scope of their training data, turning away from the programmer's possibility space. These situations may occur due to a biased dataset, or a prompt that transcends the training data.[6]

This possibility of moving past a certain possibility space resonates with Boden's transformational creativity. It thus refutes Lovelace's statement that computers are bound to what they are programmed for by introducing a surprise element to their output.

The idea of hallucination amplifies the argument. Considering the example of image creation using AI: by combining different elements of a certain training dataset and delivering an output that extrapolates the initial intention, even resulting in an unpredictable outcome, which resonates with Runco and Jaeger's (2012) notion of originality. Their notion of effectiveness can also be satisfied once an input is revealed to be useful and appropriate by itself or becomes a valuable and appropriate inspiration within the creative process. Let's suppose the hallucination does not fulfil both creativity pillars. In that case, it allows at least the detection of possible faults in the system or even influences the programmer/designer to choose a path that will better suit their needs.

5 From https://www.ibm.com/topics/ai-hallucinations (consulted on January 11, 2024).

6 A distinction between error, glitch, bug, and hallucination has to be made. To this end, the Oxford Learner's Dictionaries website was consulted on July 19th, 2024 (https://www.oxfordlearnersdictionaries.com/). An error is a misstep that can cause problems later on and affect the results. In this context, an error can be understood as a programming mistake or a bias in the dataset. A bug is similar to an error, but only applicable to a computer system or a program. An example can be a chatbot that keeps asking the same thing, without changing the topic of conversation. A glitch is a small-scale problem in which something stops working as it was supposed to. An example can be a generated image of a face with a slight distortion.

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

Applying the above typology to this context, Factual Mirages constitute a potential creative input by engaging with the prompt and further exploiting it into a deeper and more complex ideation process. A Factual Mirage in the context of the artistic, creative process can be both original and effective, as it stimulates the process by integrating new inspirational input or accelerating the process of conceptualisation by materialising different ideas (either from within or outside the scope of the possibility space) into a prototype or a draft.

4.2. Are we creative loners?

In the same event, Professor Penousal Machado points to another critique of computational creativity, which defines it as mimicry of human creativity. And here lies a problem with the conceptualisation of computational creativity. Machado states that such creativity aims for "the development of intelligent systems capable of overcoming our faults, increasing our intelligence and solving problems that otherwise we couldn't" (Beira et al., 2023). He even adds an important conception that directly refers to Runco's (2023) notion of artificial creativity. Machado states that this discussion is not so much about AI "just for the sake of it" as it is about AI and its relation to human beings and their problems. However, creativity tends to be conceptualised in a human-like manner, which poses a problem when this notion is applied to AI, considering its "increasing agency, autonomy, and creativity" (Carvalhais and Lee 2022, 72). The authors think that computational creativity is "difficult to pin down", but accessible through computational art, where computers are used as tools or as the primary medium (72-73).

In this sense, computational art decentralises the question of AI and computational creativity from anthropocentrism. By introducing computers into the creative process, the premise that only humans can be creative becomes obsolete, as they, side-by-side with a human agent or by themselves, can be creative. Humans may activate computers, and their agency may be predicted through human programming. Still, the actual computational process and the result (including the computational process) depend on the computer's agency. If we include AI and its programming capabilities in the equation, we can once again withdraw human intervention from the process. Leonel Moura's painting, acting, or poet robots are proof of this position: once triggered, they became capable of creating on their own terms, following their own concept of creativity. Hallucinations are likewise a manifestation of computational creativity, as they move beyond human intention toward a machinic expression:

Whenever there is a computational substrate, even in cases where the goal is to preserve information, there is also the potential for the computational to express itself (through glitches, bugs, or other means), and to disrupt the intended stability of the information (Carvalhais and Lee 2022, 76).

In the context of human-AI interaction, this expression opens paths to new and unpredictable sources of inspiration, as AI tools have been proven to do in normal circumstances anyway (Yannakakis et al. 2014).

4.3. Newness under the sun

History has taught us that creativity is no longer a divine trait (Glăveanu and Kaufman 2019). Humans claimed this ability for themselves, i.e., to bring something into existence, proving that, after all, there is a lot "under the sun"[7]. Modernism pulled the trigger on the idea of novelty regarding human creation. Later, with the rise of capitalism, value and profitability became another pillar of creativity (*ibid*, 17-19).

Computational creation shouldn't fall far from this. One of the biggest misconceptions regarding AI has to do with some form of human exceptionalism and superiority in relation to everything else, which leads to an understanding of computational tools as mere utensils (Moura 2023). As stated earlier, computers have their agency once triggered and can create by themselves. That includes the fact that machines can also bring forth novelty (and potentially produce value) by themselves. One can look no further than the definition of Generative AI on McKinsey's website:

> Generative artificial intelligence (AI) describes algorithms (such as ChatGPT) that can be used to create new content, including audio, code, images, text, simulations, and videos. (McKinsey & Company, 2023)

This definition explicitly attributes creative potential to computers. These entities can produce novelty, present agency, and go beyond a specific possibility space. But the question is whether computational creations are indeed new.

The concept of novelty can be questioned. The earlier mentioned observation by Leonel Moura that everything is derivative from past references can be applied. It is possible to think that everything has been done and created at some point. But other values are in play during the creative process. Creativity can be considered a social process, the social environment being of paramount importance (Glăveanu and Kaufman 2019, 17). The concept of novelty by itself addresses the creative process only on an intellectual dimension, and to create does not always imply thinking about it, but also acting about it. However, for a long time, creativity studies have preferred ideation over action. In AI and computational creativity, action can be attributed to a human agent or to Robotics if we consider a more independent computational agent.

7 From the Bible: "What has been will be again, what has been done will be done again; there is nothing new under the sun" (Ecclesiastes 1, 9), originally mentioned by Glăveanu and Kaufman, 2019.

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

5. Conclusions

The discussion on computational creativity is still far from reaching a conclusion. As literature displays two antagonistic views on the existence of computational creativity, a general consideration should be made for each one.

On one side, the argument against computational creativity displays a somewhat Orwellian reaction to this (not so) new way of working with machines. Hans Moravec is one of many voices that rang the alarm on the replacement of humans by machines, with the latter becoming the dominant form of life on the planet (Hayles 1999, 235-236). This type of reaction has been part of the history of new media. The clash of Photography and Painting is such a case (Benjamin 2008, Agüera y Arcas 2017, 16). However, nowadays there is something more at stake: the prevalence of the human being in the context of the advent of the post-human. Questions about the possible implications of the post-human involve the perpetuation of free will and human agency, or even if the liberal subject will continue to be recognised or will be overpowered (Hayles 1999, 281). The idea of the post-human as a merge between flesh and silicon, or the mere coexistence between other intelligent (and computational) entities may lead to reluctance based on anthropocentric fears that humans are becoming an inferior species and thus will be dominated, which is a scenario that science fiction has been exploring on the course of the past few decades. Scenarios like these may be responsible for a large unacceptance of computational creativity, which may be perceived as a threat.

On the other side, the visions in favour of computational creativity, namely Margaret Boden's, seem to sometimes practice a certain form of relativism due to the perspective that, in the absence of a full understanding of human and technological phenomena, we are free to make several direct connections between them. At the same time, these assumptions still carry some validity: just because we don't completely understand certain phenomena, that doesn't mean we should negate the importance of bringing machines closer to human creativity.

Returning to the question: *can a hallucination be creative,* and *can it be an input in the creative process?* A few considerations must be made beforehand.

Hallucinations, lack intentionality by themselves, although they defy the artist/programmer's intentions. It is a grey zone when it comes to the attribution of creativity. The creation process may be ignited by the human agent. However, if we accept and introduce them into the creative process, hallucinations can indeed function as a creative input because by bringing novelty, originality, and, ultimately, value they trigger the human's lateral thinking and diagrammatic reasoning (Yannakakis et al. 2014).

Additionally, on the matter of hallucinations' creative potential, they must be addressed in the context of the critique of bounded knowledge. AI tools are indeed dependent on their training dataset information. However, Leonel Moura's point that we are derivative reveals itself useful as a response: human beings depend on the information we get from our life experience, from social, cultural, economic environments, and from education (to mention a few). Also, Moura (2003) points out that the amount of information

datasets contain is much broader than the one a human brain does. Furthermore, the way AI organises and makes sense of its database differs from humans. Machinic memory is based on objective procedures, more precise storage and a better capability of bringing forth archives, meaning, a machine does not forget the same way as we humans do. Just because computers don't draw memories based on personal experiences, that does not mean that they can't bring creative inputs into existence. By identifying (novel) patterns in data instead and deviating from the initial intentions of the programmer, machines expand the notion of creativity, liberating it from human standards.

In this context, AI hallucinations can be considered creative not just when they prove useful and valuable to human creators by enhancing their creativity, but also when they independently produce a body of work that can be deemed creative solely relying on computer-driven processes. They can additionally offer insights and valuable input for the programmer/designer that are not restricted to a creative dimension – e.g., regarding questions of development or operativity.

In the realm of Arts, AI and Gen. AI specifically bring about new forms of work, both as a tool, a partner, or independently. However, there are some problems with these technologies that must be considered. To mention a few: the quality of trained data (Bender et al. 2021), which evokes issues of cultural, racial, and stylistic bias, and authorship (copyright laws, authorship legitimacy). The ways AI and Humans can create together have been studied in the last few years (for example Yannakakis et al. 2014; Figoli et al. 2022). These studies have proven that AI tools can stimulate a person's creative capacities by expanding the initial possibility space and introducing an element of surprise into the creative journey (Yannakakis et al. 2014).

Creative-wise, hallucinations are not necessarily flaws. Faults and glitches are part of creative progress, as they can disrupt an ongoing process by bringing new and unexpected inputs or deviations from the original path.

The consideration of computational creativity poses a few questions concerning the role of the (human) artist. Broeckmann (2019) states that the artist's status has been questioned since the first emergence of movements like Dadaism and Surrealism. Leonel Moura provides a possible answer. He proposes to rethink the concept of the artist. More than a doer, he is a trigger: "We [artists] trigger a process that leads to the development of something" that may be considered a work of art (2023). The artist loses control over the imagination process, which is handed to the machine. In this sense, the artist only provides the necessary input for the machine to imagine and create. Although this position may be problematic and questionable, its importance lies in how it signals a shift in the mindset regarding the place of humans in the computational creative process. Humans are part of a broader ecology composed not only of fellow organic beings but also of new machinic neighbours. Hayles presents different ways in which humans and computers may inhabit the planet. Machines may assume tasks and choices that free the human attention toward other (and probably more important) matters (Lanham *apud* Hayles 1999, 287). Another option involves placing the human element on top

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198

of a computational ecosystem that makes accurate decisions beyond human attention (Ostman and Weizenbaum *apud* Hayles 1999, 287-288). In sum, the post-human frees humans from a liberal perspective of conquering and controlling resources, moving us closer to partnerships with intelligent machines and new and better futures. "Just as the posthuman need not be antihuman, so it also need not be apocalyptic" (288).

AI Hallucinations are just one manifestation of machinic creativity. These allow us to defy computational determinism and, regarding the artistic domain, provide new gateways for successful creation and fruitful discussions on the future of the Arts. Post-humanist thought may lead to the dissolution and deconstruction of the binomial hiatus that was established between human and computational creativity by merging the two into new ways of working, raising new questions, and further expanding possibilities and capabilities: "human functionality expands because the parameters of the cognitive system it inhabits expand" (Hayles 1999, 290-291).

6. Next Steps in the Investigation

Following this article, further investigation may involve field research alongside designers to understand how AI hallucinations inform the creative process; and how the theoretical proposal of this paper finds its concretisation. A more complete investigation could trace a framework that applies to areas such as Fashion Design, Graphic Design, and Industrial Design, among others.

Another possible way for future investigation involves authorial rights and authorial legitimacy. When artworks have been created with AI, authorship may be divided between several actors, namely the artist who triggered the AI agent, the AI agent itself, the database collector, or even the authors of the data in the dataset.

A third proposal for the extension of this inquiry highlights the concept of creativity in non-human terms. This paper focuses on computational creativity mainly from a perspective that underlines the role of technological companionship in relation to a human agent. Developing more studies on computational creativity and its application in the field could be insightful. A deeper reflection on past and more recent work developed by Katherine Hayles and Donna Haraway, alongside the advancements in the field of Artificial Life, will provide the ground for new perspectives on computational creativity.

—

References

Agüera y Arcas, Blaise. 2017. "Art in the Age of Machine Intelligence." *Arts* 6 (4): 18. https://doi.org/10.3390/arts6040018.

Associação para a Promoção e o Desenvolvimento da Sociedade de Informação. 2023. "Conferência Inteligência Artificial e a Arte." YouTube vídeo, 5 de novembro de 2023, 06:19. https://www.youtube.com/watch?v=AodJ7Wiy2uI.

Beira, João, Cardoso, Inês, Guerra, Ana, & Machado, Penousal. 2023. *Pode a criatividade ser não humana?* Conference Session. Inteligência Artificial e a Arte, Lisbon, Portugal. https://www.youtube.com/watch?v=AodJ7Wiy2uI

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? '. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. Virtual Event Canada: ACM. https://doi.org/10.1145/3442188.3445922.

Benjamin, Walter. 2008. *The Work of Art in the Age of Mechanical Reproduction*. London: Penguin Books.

Boden, Margaret A. 2004. *The Creative Mind: Myths and Mechanisms*. 2nd ed. London ; New York: Routledge.

Broeckmann, Andreas. 2019. "The Machine as Artist as Myth." *Arts* 8 (1): 25. https://doi.org/10.3390/arts8010025.

Carvalhais, Miguel, and Rosemary Lee. 2022. "Spectral and Procedural Creativity: A Perspective from Computational Art." *Transformations Journal* 36: 71–81. http://www.transformationsjournal.org/2022-issue-no-36-artificial-creativity/.

Crawford, Kate, and Trevor Paglen. 2021. "Excavating AI: The Politics of Images in Machine Learning Training Sets." *AI & SOCIETY* 36: 1105–1116. https://doi.org/10.1007/s00146-021-01162-8.

Curry Jansen, Sue. 2021. *What Was Artificial Intelligence?* 1st ed. USA: mediastudies.press. https://assets.pubpub.org/52h6z8jg/61648745415960.pdf.

Del Campo, Matias, Alexandra Carlson, and Sandra Manninger. 2021. "Towards Hallucinating Machines - Designing with Computational Vision." *International Journal of Architectural Computing* 19 (1): 88–103. https://doi.org/10.1177/1478077120963366.

Figoli, Fabio Antonio, Lucia Rampino, and Francesca Mattioli. 2022. "AI in Design Idea Development: A Workshop on Creativity and Human-AI Collaboration." In *DRS2022: Bilbao,* edited by D. Lockton, S. Lenzi, P. Hekkert, A. Oak, J. Sádaba and P. Lloyd, 25 June - 3 July, Bilbao, Spain. https://doi.org/10.21606/drs.2022.414.

Galanter, Philip 2009. "Thoughts on Computational Creativity." In Computational Creativity: An Interdisciplinary Approach. Dagstuhl Seminar Proceedings, vol. 9291, Schloss Dagstuhl – Leibniz-Zentrum für Informatik. https://doi.org/10.4230/DAGSEMPROC.09291.32.

Glăveanu, Vlad P., and James C. Kaufman. 2019. 'Creativity: A Historical Perspective'. In *The Cambridge Handbook of Creativity*, 2nd ed., edited by James C. Kaufman and Robert J. Sternberg, 11–26. Cambridge University Press. https://doi.org/10.1017/9781316979839.003.

Hayles, N. Katherine. 1999. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics.* Chicago: University of Chicago Press.

IBM. "AI Hallucinations." Accessed January 11, 2024. https://www.ibm.com/topics/ai-hallucinations.

Jefferson, G. 1949. "The Mind of Mechanical Man." *BMJ* 1 (4616): 1105–10. https://doi.org/10.1136/bmj.1.4616.1105.

Mitchell, Melanie. 2021. "Why AI Is Harder Than We Think." ArXiv. http://arxiv.org/abs/2104.12871.

Nake, Frieder. 2016. 'The Disappearing Masterpiece'. In *Proceedings of the Fourth Conference on Computation, Communication, Aesthetics, and X*, 11–26. Bergamo.

Pasquinelli, Matteo, and Vladan Joler. 2021. "The Nooscope Manifested: AI as Instrument of Knowledge Extractivism." *Ai & Society* 36 (4): 1263–80. https://doi.org/10.1007/s00146-020-01097-6.

Rawte, Vipula, Swagata Chakraborty, Agnibh Pathak, Anubhav Sarkar, S.M Towhidul Islam Tonmoy, Aman Chadha, Amit Sheth, and Amitava Das. 2023. "The Troubling Emergence of Hallucination in Large Language Models - An Extensive Definition, Quantification, and Prescriptive Remediations." In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 2541–73. Singapore: Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.155.

Runco, Mark A. 2023. "AI Can Only Produce Artificial Creativity." *Journal of Creativity* 33 (3): 100063. https://doi.org/10.1016/j.yjoc.2023.100063.
Runco, Mark A., and Garrett J. Jaeger. 2012. "The Standard Definition of Creativity." *Creativity Research Journal* 24 (1): 92–96. https://doi.org/10.1080/10400419.2012.650092.
Searle, John R. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (3): 417–24. https://doi.org/10.1017/S0140525X00005756.
Turing, A. M. 1950. "I.—Computing Machinery and Intelligence." *Mind* 59 (236): 433–60. https://doi.org/10.1093/mind/LIX.236.433.
____. 1996. "Intelligent Machinery, A Heretical Theory." *Philosophia Mathematica* 4 (3): 256–60. https://doi.org/10.1093/philmat/4.3.256.
Turing, Alan, D. Ince, and Alan Turing. 1992. *Mechanical Intelligence*. Collected Works of A.M. Turing. Amsterdam; New York, NY: North-Holland.
Weiner, Robert Paul. 2000. *Creativity & beyond: Cultures, Values and Change*. Albany: State University of New York Press.
Yannakakis, Georgios N., Julia Liapis, and Constantine Alexopoulos. 2016. "Mixed-Initiative Co-Creativity." *Proceedings of the 2016 Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*. Accessed January 11, 2024. https://www.semanticscholar.org/paper/Mixed-initiative-co-creativity-Yannakakis-Liapis/0a410eecf3b23042f95ebfbe0ffdc08ea697c9d9.

Biographical note
João G. Patrício is a researcher currently pursuing a PhD in Digital Media at Universidade Nova de Lisboa, focusing on the intersection of fashion design, generative artificial intelligence, and creativity studies. He holds a Master's in Digital Marketing from Universidad de Salamanca and IME Business School, and a Master's in Communication, Culture, and Information Technology from ISCTE. With professional expertise in SEO and Creative Industries, he explores the evolving relationship between technology and media. João has authored works on SEO and Semantic Web in the context of online journalism and served as an invited professor at Universidad de Salamanca.

ORCID
0009-0008-2806-1119

CIÊNCIA ID
CF14-7E76-3F3B

Institutional address
ICNOVA — Instituto de Comunicação da NOVA
Universidade Nova de Lisboa — Faculdade de Ciências Sociais e Humanas NOVA FCSH
Av. de Berna, 26 C, 1069-061 Lisboa, Portugal

Declaration of conflicting interests
The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

To cite this article
Patrício, João G. 2024. "Is Delulu the New Trululu? Artificial Intelligence Hallucinations as Input in the Creative Process." *Revista de Comunicação e Linguagens* (60-61): 226-245. https://doi.org/10.34619/jwrp-nbht.

Received Recebido: 2024-02-26 **Accepted** Aceite: 2024-04-06

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.60-61 (2024) ISSN 2183-7198