Generative artificial intelligence as humans' fourth narcissistic wound

A inteligência artificial generativa como quarta ferida narcísica do humano

LUCIA SANTAELLA[a]
Pontifical University of São Paulo. São Paulo – SP. Brazil.

DORA KAUFMAN[b]
Pontifical University of São Paulo. São Paulo – SP. Brazil.

ABSTRACT

The progress of artificial intelligence, particularly with generative AI models, has provoked intense reactions, regardless of whether they are based on the logic and functioning of the technology. Unlike predictive AI, generative AI produces original content by synthesizing texts, images, voices, videos, and code from large databases and may significantly impact the creative economy. This study introduces the basic concepts of AI and generative AI (including a taxonomy of generative models) and outlines the distinction between image or video and text production techniques. The central argument of this study claims that the cultural fuss is not accidental, defending the hypothesis that the advent of generative AI places humanity amidst the crossing of its fourth narcissistic wound.

Keywords: AI, generative AI, ChatGPT, Freud, narcissistic wound.

RESUMO

O avanço da inteligência artificial (IA), particularmente com os modelos de IA generativa, tem provocado intensas reações, fundamentadas ou não na lógica e no funcionamento da tecnologia. Distinta da inteligência artificial preditiva, a IA generativa produz conteúdo original sintetizando texto, imagem, voz, vídeo e códigos a partir de grandes bases de dados, com potencial de impactar significativamente a economia criativa. Este artigo introduz conceitos básicos da IA e a generativa, incluindo uma taxonomia dos modelos generativos, e delimita a distinção entre as técnicas de produção de imagem ou vídeo e as de produção de textos. O argumento central deste artigo é que o alarido cultural não é casual, defendendo-se a hipótese de que o advento da IA generativa coloca a humanidade em plena travessia de sua quarta ferida narcísica.

Palavras-chave: IA, IA generativa, ChatGPT, Freud, ferida narcísica

[a] Full professor at the Pontifical Catholic University of São Paulo (PUC-SP) in the Technology, Intelligence, and Digital Design (TIDD) and Communication and Semiotics Programs. Ph.D. in Literary Theory at PUC-SP and Professorship Habilitation in Communication Sciences at the School of Communication and Arts of the University of São Paulo (ECA/USP). Researcher 1A at CNPq. Author of 51 books and organizer of 26 books. Orcid: https://orcid.org/0000-0002-0681-6073. E-mail: lbraga@pucsp.br

[b] Professor of the Technology, Intelligence, and Digital Design Program at the Pontifical University of São Paulo (TIDD/PUC-SP). Ph.D. from the School of Communication and Arts of the University of São Paulo (ECA/USP) with an International Ph.D. Internship at the Université Paris Sorbonne IV, France, Post-doctorate at the Federal University of Rio de Janeiro (UFRJ) and TIDD/PUC-SP. Author of the books *A Inteligência Artificial Irá Suplantar a Inteligência Humana?* and *Desmistificando a Inteligência Artificial.* Orcid: https://orcid.org/0000-0001-7060-4887. E-mail: dkaufman@pucsp.br

DOI: http://dx.doi.org/10.11606/issn.1982-8160.v18i1p37-53

Generative artificial intelligence as humans' fourth narcissistic wound

ENABLING MACHINES TO solve tasks based on logical reasoning, which humans often perform intuitively, was a challenge in the early days of artificial intelligence (AI), a field of research that developed since the middle of the last century. Several attempts involving formal languages supported by logical inference rules (symbolic AI) have had limited success, suggesting the need for systems to generate their own knowledge by extracting patterns in data, i.e., to "learn" from data without receiving explicit instructions. This process is usually called "machine learning," a subfield of AI that today is undoubtedly the largest field of AI in terms of the number of practitioners (Alpaydin, 2016; Bengio et al., 2016; Kaufman, 2022b).

Multiple observable or unobservable factors influence the learning process of these systems in the physical world, which is subject to effects from external sources. For example, the pixels in an image of a red car may be very close to the color black at night, and the shape of a car's silhouette varies with the viewing angle. The machine learning technique that best solves these challenges today is deep learning (deep learning neural networks – DLNNs). Inspired by the functioning of neural networks in the biological brain, it is capable of dealing with high-dimensional data (for example, millions of pixels in an image) by introducing complex representations, expressed in terms of other, more straightforward representations organized into several layers, presenting positive results in several areas, particularly in computer vision, voice, and image recognition. Additionally, DLNNs establish correlations that are not perceptible to humans, whose tendency is to consider only the "strongest" correlations. However, when grouped, the "weakest" ones can significantly impact the models' accuracy (Kaufman, 2022a).

Despite notable advances, AI still lacks a unifying theory covering the foundations for creating "intelligent machines." What we have are empirical models (Kouw et al., 2013). In any case, AI has evolved, and the degree of permeation of its applications in society and human life has jointly grown. With the advent of generative AI models, the point at which AI finds itself today has caused an absolute cultural uproar. It is not accidental, given that its applications, although nothing more than simulations, come very close, or rather, they convincingly imitate human skills. Thus, this article aims to lay the groundwork to launch and defend the hypothesis that the advent of generative AI is placing humanity in the midst of crossing its fourth narcissistic wound.

GENERATIVE ARTIFICIAL INTELLIGENCE

2022 was a particularly significant year for generative AI: Google's Imagen launched in May; Stability.AI, from Stable Diffusion, in August; and OpenAI launched DALL-E in July, DALL-E 2 in September, and ChatGPT in November. One of the indicators of expectations regarding the transformative power of generative AI is the reaction of investors: startups driven by generative AI received enormous contributions from investment funds, reaching astronomical market values for companies in their infancy, as in the case of Jasper, with an estimated market value of US$ 1.5 billion and raised US$ 125 million in October; and Stability. AI, which, valued at US$1 billion, raised US$ 101 million in the same month. Throughout 2022, investors pumped at least US$ 1.37 billion into 78 generative AI businesses, almost the same amount invested over the past five years. Currently, 450 generative AI startups are estimated to exist (Griffith & Metz, 2023).

Generative AI, distinct from predictive machine learning models (predictive AI)—focused on extracting patterns from data and making predictions in specific tasks—produces original content from large databases, i.e., it uses data to generate more data, synthesizing text, images, voice, video, and codes. These models have the potential to impact the so-called creative economy with practical applications significantly: a) in the automated generation of content in articles, blog posts, and social media; b) in boosting the quality of content due to the training of its algorithms, and the use of large databases to identify patterns that surpass human cognition; c) in the production of more diverse content, including text, images, and video; and d) generating personalized content based on users' profiles and preferences (Davenport & Mittal, 2022). Jasper, for example, is being applied in marketing actions to produce blogs, social network posts, web texts, sales e-mail messages, and advertisements, among other content for interaction with users, clients, and consumers. DALL-E 2, aimed at generating images, is being applied in producing advertising pieces for leading companies in the segment (Davenport & Mittal, 2022).

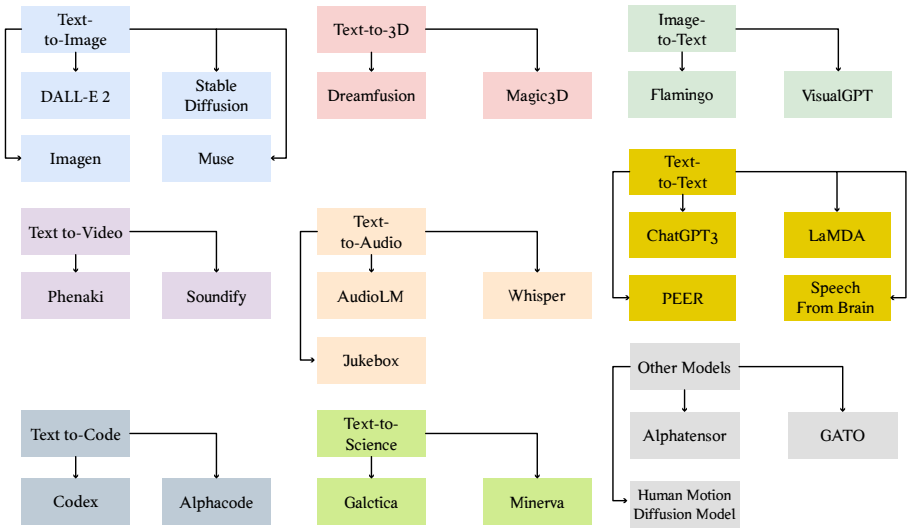
Generative AI models derive from different architectures of the deep neural network technique. "Architecture," in this case, represents how the components of neural networks—artificial neurons, layers, and connections—are organized. The first of these architectures was generative adversarial networks (GAN), proposed by Bengio et al. (2014), with positive results in the health area—such as generating synthetic data and improving computed tomography or magnetic resonance images, reducing the time patients are exposed to radiation—but equally with negative results when generating deep fakes.

Currently, the most popular generative AI solution is ChatGPT, made available for public experimentation by OpenAI on November 30, 2022, based on the Transformer architecture developed by the Google Brain team (2017). Initially for word translation processes, Transformer became the preferred architecture for natural language processing (NLP) models, being, for example, the basis of the OpenAI series, GPT-base, GPT -2, GPT-3, GPT3.5-turbo, GPT-4, and GPT-4-turbo, which in turn are the basis of ChatGPT. The differentiating characteristic of Transformer is that it is trained in dialogues, making it possible to capture nuances, distinguish the fluidity of a conversation, and generate responses that apparently make sense (Uszkoreit, 2017).

Garrido-Merchán and Gozalo-Brizuela (2023) organized the generative models into a taxonomy, resulting in nine categories, represented in Figure 1.

Figure 1

Taxonomy of the most popular generative AI models



Note. Garrido-Merchán & Gozalo-Brizuela (2023, p. 3).

These systems require robust computing capacity, professional expertise, and access to large volumes of quality data, implying high model development and implementation levels. This restricts access to a limited number of companies, favoring the already expected market concentration. To illustrate, generative models can handle data from all Wikipedia, all social networks, or all images from Google Search. Figure 2 lists the most popular models released in the last two years.

Figure 2

Generative AI model categories

Category	Description	Representative Models
Text-to-image	Input text prompt, output is an image	DALL-E 2, IMAGEN, Stable Diffusion, Muse
Text-to-3D	Text input, 3D image output, special for games	Dreamfusion, Magic3D
Image-to-Text	Image input, text output, recommended for describing an image	Flamingo, VisualGPT
Text-to-Video	Text input, video output (image sequence)	Phenaki, Soundify
Text-to-Audio	Text input, audio output, critical for videos, music, and other contexts	AudioLM, Jukebox, Whisper
Text-to-Text	Text input, unpublished text output, suitable for Q&A	ChatGPT. LaMDA, PEER. Meta AI Speech from Brain
Text-to-Code	Text input, code output, special for programming	Codex, Alphacode
Text-to-Science	Text input, output scientific article (very embryonic)	Galactica, Minerva

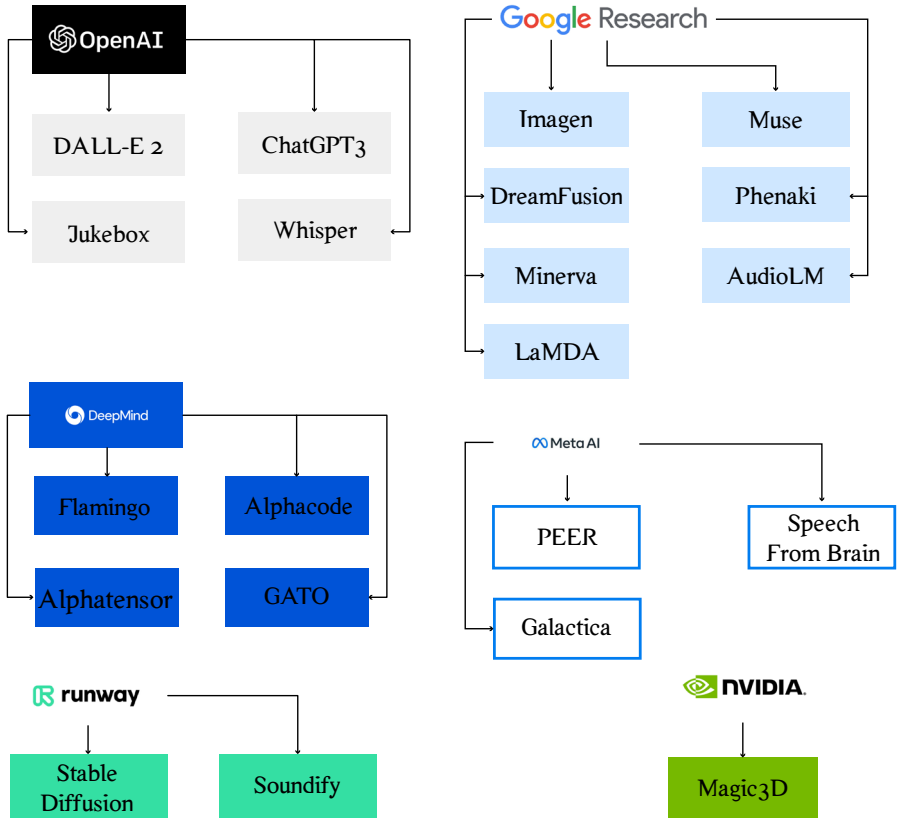
Note. The authors elaborate on the taxonomy by Garrido-Merchán & Gozalo-Brizuela (2023).

Figure 3 shows the same models classified by the developer, accounting for only six organizations (Garrido-Merchán & Gozalo-Brizuela, 2023).

These applications require the involvement of human experts throughout the development and implementation process, and the user of the models will necessarily need to test several instructions to obtain the desired result. Then, a human being must evaluate and edit the generated content. In the case of models that generate image content, the synthetic images (generated by AI) must be manipulated by a human expert. Jason Allen, the winner of Colorado's digitally manipulated photography contest using the Midjourney system, took over 80 hours and more than 900 versions to perfect the image's quality and sharpness with Adobe Photoshop and other AI resources (Vincent, 2022).

Figure 3

Models grouped by developer



Note. Garrido-Merchán & Gozalo-Brizuela (2023, p. 4).

In 2020, before the launch of ChatGPT, Massimo Chiriatti and Luciano Floridi (2019) identified the limitations of GPT-3 through three tests based on mathematical, semantic, and ethical issues. In the mathematics test, the GPT-3 satisfactorily performed when asked to perform simpler operations but was disappointing when some zeros were added, i.e., for larger numbers. In the semantic test, as GPT-3 has no understanding at all, only the statistical ability to associate words, it could not answer the question, "Tell me how many feet fit in a shoe?" In the ethics test, GPT -3 was trained on databases generated by human beings, and thus, it reproduced the prejudices of human society, such as gender and ethnicity.

ChatGPT was featured in the "MONTREAL.AI Debates Series," an event organized by Québec Artificial Intelligence (Québec.AI). Throughout the panels,

participants illustrated flaws, serious errors in simple logical problems, and false and inconsistent statements, highlighting the absence of four aspects of human cognition: a) abstraction, b) reasoning, c) compositionality (does not understand language in terms of a whole composed of parts, like human beings); and d) factuality (their updates are not incremental from new facts, thus requiring retraining).

However, the apparent consistency of ChatGPT's responses leads the user to mistake them for accurate and true. The enchantment, or even magic, of this unprecedented interface in dialogue format needs to be relativized, avoiding the hype that makes it challenging to identify the tangible benefits; and, on the contrary, we must seek to mitigate their potential harm, such as the even greater spread of misinformation and fake news (the creation of deep fakes, for example, until now required specialized skills from their developers, while AI models generative techniques give broad access to non-specialists). OpenAI has warned that ChatGPT may occasionally generate incorrect information and produce harmful instructions or biased content and that this research project will continue to be refined. In other words, precision (or lack thereof) is still a problem that needs to be overcome by generative models (Kaufman, 2023).

Evaluative reflections

Considering user adoption, it is worth highlighting the fundamental difference between image or video production techniques and text production techniques. In the first case, the four applications (Imagen, Stability.AI's Stable Diffusion, DALL-E, and DALL-E 2) had a significant impact on the world of image production, both for artists and visual designers and producers in general, since, in any of these models, commands activated in a relatively simple interface are enough to generate visual images, including videos, shaped according to the terms of the statement. The impact is not accidental, as it is a significant process of automatic intersemiotic translation that, through specific neural networks, transposes the textual code into the image code. During 2022, the impact of these launches only rehearsed the immense repercussions that ChatGPT has significantly caused.

ChatGPT's surpassing power of repercussion, concerning other applications of generative AI, is explainable because, while the production of images and videos affects only a niche of human creation—occupied by professionals who operate in the innumerable fields of visuality—ChatGPT affects any human being in their linguistic ability. This particularity helps us understand why the fuss over imaging systems was not as deafening, nor did it awaken the same socio-cultural reverberations as ChatGPT, despite there being considerable

protests from artists and designers against the use of unauthorized copies of their works to feed the databases (Gagliardi, 2022; Butterick, 2023).

While the texts that result as ChatGPT responses to user commands are impressive, to achieve its results, according to Wolfran (2023), what the chatbot seeks to do is produce a "reasonable continuation" of any text it has obtained at the time. By "reasonable," one should understand "what one might expect someone to write after seeing what people have written on billions of webpages." In fact, the model accesses billions of pages of text written by humans and then estimates the probability of the word that best fits the text sequence to meet the user's request. Like every statistical probability system, it intrinsically has an uncertainty variable responsible for part of the imperfections.

So, when writing an essay, the chatbot is just repeatedly asking itself, "Given the text so far, what should the next word be?" and, with each answer, add a new word. In linguistics terminology, what is added is a token, which can be just a part of a word. Therefore, sometimes, the chatbot can even invent a new word. Overall, what it gets is a list of words with probabilities. Which one does it choose? The best classified, i.e., the one with the highest parameter—weight assigned to the tokens in the training process, in which, through correlations, the system's algorithms establish a hierarchy between the tokens. Thus, it produces flat texts. However, because of randomness, if the user uses the same command several times, different essays will likely be obtained each time (Wolfran, 2023).

Models are built to persuade. In other words, they have been trained to convince human beings that they carry on a conversation as if they were human. To do this, they can even lead us to believe that they have emotions and are capable of feeling. Therefore, they are on the path to becoming "friends to the lonely and assistants to the tormented," among many other activities, to the point that generative AI fuels fears that their models will be able to replace, without mediation, the work of dozens of writers, graphic designers, and form fillers, among others (Klein, 2023). In short, instead of appearing as alien, in the differences it maintains with humans, generative AI, on the contrary, appears as too human. This appears to be because, in ChatGPT, the texts produced are syntactically consistent and semantically coherent. The grammatical correction is almost perfect, considering an adequate lexical selection supports it, and impeccable phrasal contiguity exists. To this extent, having discarded some contextual errors, often absurd and laughable—but evident in a prison system of language that lacks the common sense of lived life—what remains to be confirmed is the level of specialization of the textual content, which can only be checked by a specialist. For an average receiver, therefore, although always very basic, the texts produced are riskily convincing.

The recent cases of generative AI models of fake images and videos are as dangerous or even more dangerous. The technically almost perfect images, with visual noise imperceptible to unsuspecting eyes, of the Pope wearing a puffer coat, Trump in prison, and King Charles wearing a Russian hat, for example, fall into the category of deep fakes, with all the harmful consequences brought about when the vision and understanding of reality are wrong. According to C. S. Peirce's theory of perception (Santaella, 2021), we are not immediately capable of doubting what we see. This occurs because vision is indissolubly accompanied by a perceptual judgment that informs us about what is seen. We cannot doubt this, or living would be impossible. We can only doubt what appears to our eyes when, for some reason, we are led to subject perception to critical scrutiny caused by suspicion. This dangerously means that the banality of evil (Arendt, 1999) and human perversity in action can fuel the fake news industry in volumes far more significant than those already known.

In any case, critical alerts and concerns about the necessary and urgent regulation of AI cannot lead to the erasure or minimization of the fact that, following the words of Huttenlocher *et al.* (2023), generative AI presents a philosophical and practical challenge on a scale not experienced since the beginning of the Enlightenment, to the extent that, with all the risks and surprises, it will open revolutionary paths for human reason and new horizons for consolidated knowledge. For the authors, however, there are categorical differences concerning the Enlightenment, in which knowledge was achieved progressively, step by step, with each step testable and teachable. "AI-enabled systems start at the other end. They can store and distill a huge amount of existing information[…] billions of items. Holding that volume of information and distilling it is beyond human capacity" (Huttenlocher *et al.*, 2023).

While the authors' comparison is relatively legitimate, the differences from the Enlightenment are much more categorical and profound and go beyond aspects related to rationality and human knowledge. After Kant's 18th century, belief in reason experienced successive setbacks. First, remember Goya's well-known statement that "the sleep of reason produces monsters." The history of the 20th century is enough to demonstrate that, when awakened, reason is also and perhaps even more capable of producing monsters. In the 19th century, Marx demonstrated that capitalism was storming traditional values, dismantling everything solid into thin air (Berman, 1983). Then, at the beginning of the 20th century, Max Weber (1967) pointed to the disenchantment of rationalism in world domination. Meanwhile, Freud demolished, among other human illusions, the cult of free will and autonomy. There is also the Nietzschean blow

to the security of the divine. More recently, diagnoses led Bauman (2001) to the analysis of liquid societies and Sloterdijk (2006) to the instabilities of foams.

In this context of uncertainty, ambivalence, contradictions, and unpredictability, AI advances have been taking place, reaching ever closer to the supposed human supremacy. Predictive AI algorithms monitor and manipulate our lives in relatively invisible ways. Their effects are felt, but the operations that control them remain outside our jurisdiction, i.e., they are not transparent. Generative AI, on the other hand, penetrated the most intimate human secret: the linguistic and semiotic potential that, until then, constituted the inimitable hallmark of *Sapiens*. The issue is so disturbing that it led Harari *et al.* (2023) to sensationally claim that "AI's new mastery of language means it can now hack and manipulate the operating system of civilization. By gaining mastery of language, AI is seizing the master key to civilization, from bank vaults to holy sepulchers."

Even if such a prognosis may prove effective in the unpredictable future, at present, the statement violates the truth of the facts by granting AI an autonomy that it does not have. Without remembering the obvious issue that ChatGPT regurgitates combinations between billions of texts produced by humans, the chatbot does not work alone, i.e., it involves being activated by human users in the form of linguistic dialogue that, until then, was the exclusive prerogative of humans. It is precisely there, in the fact that humans feel in some way equal in their key claim to exclusivity, that the socio-cultural and even psychic explosion that generative AI has been causing has been fueled. This leads us to propose and defend the hypothesis that, with AI at its point today, humanity is in the midst of crossing its fourth narcissistic wound.

THE FOURTH NARCISSISTIC WOUND

At the end of 1916, a Hungarian editor invited Freud to contribute an article to the journal *Nyugat*. Three months later, the article was published in German in *Imago*. Under the title "A Difficulty in the Path of Psycho-Analysis" (Freud, 1996a), the article became known and widely cited as "The Three Narcissistic Wounds." Freud returned to the topic on other occasions in the more general context of his discussions about resistance to psychoanalytic theories. The first part of the article summarized the text "Introduction to Narcissism" (Freud, 1996b), which Freud had written in 1914. The second part discussed the difficulty of psycho-analysis. The article is brief and is aimed at a lay but educated reader. According to Freud, this is important because the difficulty is not intellectual but affective and alienating. "Where sympathy is lacking, understanding will not come very easily" (Freud, 1996a, p. 85).

The premises of the discussion are found in the theory of libido: "at the beginning of the development of the individual all his libido (all his erotic tendencies, all his capacity for love) is tied to himself." Later, vital needs cause the libido to flow from the ego to external objects. There is, therefore, a progression from narcissism to object love. However, a certain amount of libido is always retained by the ego, persisting as narcissism. This very brief outline is necessary so that we can understand that narcissism does not necessarily operate in its form of neurosis. However, it is valid in other situations, such as, for example, the narcissism of children or "the excessive narcissism of primitive man that we ascribe his belief in the omnipotence of his thoughts and his consequent attempts to influence the course of events in the external world by the technique of magic" (Freud, 1996a, p. 87).

It is, therefore, a universal narcissism of human beings, the defense of their self-love, which, according to Freud, suffered three severe blows resulting from scientific research. The first was the Copernican blow. The belief in the central role of the Earth, our home—a belief also contested by Giordano Bruno who, with his cosmic pluralism, displaced the planet Earth and the human being from the center of the universe—and their role of dominance in the universe seemed to fit together very well to the human being's propensity to consider themselves the lord of the world. The destruction of this narcissistic illusion acted for humanity as a cosmological blow to their self-love.

The second blow, which hit human narcissism, was Darwinian. Human beings placed an abyss of separation between their nature and that of animals, attributing to themselves the exclusive sovereignty of reason and divine ascendancy, believing in their supreme position over other animals in the biosphere. Unlike Descartes and his followers, who considered animals to be beastly machines because they lacked a soul, it was well remembered by Freud that this arrogance is not part of the world of children nor of the primitive totemism that attributed their ancestry to an animal ancestor. The research of Darwin, his precursors, and collaborators, based on scientific data, knocked out, with a biological blow, the human presumption about their separation from the animal kingdom and forced the recognition that all living beings descend from a common ancestor.

Finally, for Freud, the third blow is perhaps the one that hurts the most. Although externally humiliated, narcissism continued to speak loudly, as the human persisted in feeling like the master of his own home, that of his mind and conscience. "Somewhere in the core of his ego he has developed an organ of observation to keep a watch on his impulses and actions and see whether they harmonize with its demands. If they do not, they are ruthlessly inhibited and withdrawn." The illusion seems convincing.

Its internal perception, consciousness, gives the ego news of all important occurrences in mental operations, and the will, directed by this information, executes what the ego orders and modifies everything that seeks to be realized spontaneously. This occurs because the mind is not a simple thing. On the contrary, it is a hierarchy of superior and subordinate instances, a labyrinth of impulses that strive towards action independently of each other, corresponding to the multiplicity of instincts and relationships with the external world, many of which are antagonistic and incompatible. Proper functioning requires the highest of these instances to be aware of everything happening and its will to penetrate everything to exert its influence. In effect, the ego feels secure regarding the integrity and reliability of the information it receives and the openness of the channels through which it imposes its orders (Freud, 1996a, p. 88).

However, psycho-analysis provides us with a version that does not match this self-control. While we feel relatively confident that we know everything in our minds, this is a big mistake. In fact, much of our own minds' activity is outside our will and access. Intelligence fails because the information we have is enigmatic, as well as incomplete, as both intelligence and consciousness are overdetermined by the unconscious, which makes us, to some extent, unknown to ourselves. At that time, Freud had not yet published Beyond the Pleasure Principle (Freud, 1996c), which would take the development of his ideas increasingly more profound into the complex effects of the unconscious. Obviously, we are not going into these details here, as what is essential for the continuity of the argument is to highlight that today, with AI, we are faced with the fourth narcissistic wound under the action of the blow dealt by the appearance of ChatGPT.

Mazlish detected the fourth discontinuity in a book less cited than it should be, *The Fourth Discontinuity: The Co-evolution of Humans and Machines* (Mazlish, 1993). The author takes up the three blows already developed by Freud but gives them a new interpretation or a reinterpretation. Mazlish borrowed the term "discontinuity," as well as "breaks" or "gaps," from Jerome Bruner (1956), who considered blows to narcissism as the elimination of false discontinuities and the consequent restoration of necessary continuities between the Earth, the solar system, and the cosmos (Copernicus/Galileo Galilei), between the human animal and other animals (Darwin), and between the human and the lack of self-knowledge (Freud). From this, Mazlish introduced the fourth discontinuity between humans and machines, demanding overcoming this unsustainable rupture despite the new blow to human self-esteem that it must cause. This is an inevitable blow, as humans and the machines they create are continuous and, therefore, inseparable.

There is much evidence that the emergence of *Sapiens* was due to the continuous interaction of technologies with their physical, emotional, and mental transformations. Much later, amid the Industrial Revolution, the continuity between humans and their technologies did not go unnoticed by Karl Marx, so much so that, for him, "Technology reveals man's dealings with nature, discloses the direct productive activities of his life, thus throwing light upon social relations and the resultant, mental conceptions." (Marx *apud* Mazlish, 1993, p. 5).

Still, in direct opposition to Cartesian and neo-Cartesian discontinuities, Pascal's work on calculating machines was taken up by people such as the eccentric 19th-century mathematician Charles Babbage, whose brilliant designs surpassed the technology available to him. Therefore, it was necessary to wait for more than a century for the combinations of mathematics, physical experimentation, and modern technology to create the machines that now confront us and reawaken the myths of discontinuity. According to Mazlish, humans feel threatened by machines and disharmony with their extensions because they establish a gap, and even an abyss, between themselves and the technologies that expand their capabilities.

While the diagnosis is raised there, what we are interested in defending, based on Mazlish, is the difference, unfortunately little remembered, between machines that extend and complement musculature, physical strength (industrial machines, portrayed and ridiculed in the film Modern Times, by Chaplin, still present in robotic factories), and those that extend human sensoriality and mental capacity, as these are the ones that led to the scripts that today lead to ChatGPT and the fuss it has caused.

Sensory machines are communicating machines. They started with the camera and the impact it had on their contemporaries. Two texts escaped the commonplace of euphoria or dysphoria and, for this very reason, became anthologies. W. Benjamin's text (1975) on the "era of technical reproducibility" due to the rupture it caused in the values of creativity cherished in the past, and V. Flusser's text (1985) on the "black box philosophy." Flusser did well when abandoning the notion of machine for his ingenious conception of apparatus. In fact, the gramophone, photography, cinema, radio, and television do not fit the idea of a machine. The technological sophistication that constitutes them is based on the fact that they have internalized the techniques and functions of our sensory organs, extending, or instead expanding, these functions beyond our bodies. McLuhan (1969) posits that the media are extensions of the human, especially television, as an extension of the central nervous system. More than that, by extending our perceptual organs, they create environments or socio-cultural and political ecologies of their own.

Generative artificial intelligence as humans' fourth narcissistic wound

Just when television was at the height of its popularity, in the 1950s and 1960s, the computer was already beginning to scale. It began in the form of enormous equipment that regurgitated punched cards, which made it possible to call it a machine, a black box whose secrets were only accessible to systems analysts and other specialists. It did not take long for it to acquire more familiar features, entering our homes under the name "personal computer" and, without much delay, transforming itself into a communicating device, a media of all media, connecting users across the planet. Unfortunately, despite all the intelligent operations within it, invisible to the user, such as protocols, with their rules and standards composing a universal language, even their interactive interfaces allowing access, sending, and exchange of messages and files, despite the hypermedia language constitutive of the networks, all of this started to be conceived merely as a resource. In this derogatory way, a hiatus was produced, a fourth type of discontinuity between human intelligence and computational intelligence, as if software, connections between hardware and software, and algorithms were simple mechanical operations incapable of calling into question the supremacy and exclusivity of human intelligence.

Such strategically separatist conditions in defense of narcissism did not cease to exist even when the smartphone reached our hands, bringing a deluge of screens populated with languages of the most different genres and species, all together and at the same time, under the command of platforms and applications with which we very quickly learn to interact. Given this, to maintain the protectionism of human self-esteem, the solution is that the tool works because it is made by humans, programmed by humans, leaving it with no agency or intelligence of its own.

However, just over ten years ago, a new visitor entered the culture scene: artificial intelligence. To begin with, the name "intelligence" seems offensive to maintaining the necessary dose of narcissism. Therefore, in favor of defending the fourth discontinuity, claims abound that the name is nothing more than a mistake that needs to be corrected, as artificial intelligence is not intelligent. Indeed, the statements do not explain how intelligence is understood (there are dozens of definitions of intelligence), maintaining as a backup the protection of the exclusivity of human intelligence. An exclusivity that is not shaken even by the fact that the algorithms are no longer the same, that they are trained to "learn," and that their performances are developed in neural networks that mimic the neural networks of our brain in layers of highly complicated statistical operations.

Even considering the lack of consensus on the necessary differentiations between the concepts of intelligence, consciousness, mind, thought, and other similar concepts, it is not difficult to detect the action of the fourth discontinuity

behind the peremptory certainty with which it is proclaimed that AI is not intelligent. However, recently, societies were shaken by the entry of ChatGPT directly into people's lives, an abrupt entry to non-specialist users cultivated in the scripts of researchers. The fuss that this new AI character is causing is such that it led us to raise the hypothesis that it definitely presents itself as a fourth blow struck against the protected human self-esteem, constituting itself as the "fourth narcissistic wound" since it inserts itself into the most distinctive and deeply human stronghold: our linguistic capacity and the skills for conversation and dialogue. The die is cast. Only time will tell how human beings will enter the game and what skills and strategies they will play.

REFERENCES

Alpaydin, E. (2016). *Machine learning*. MIT Press.

Arendt, H. (1999). *Eichmann em Jerusalém: Um relato sobre a banalidade do mal* (J. R. Siqueira, Trad.). Companhia das Letras.

Bauman, Z. (2001). *Modernidade líquida* (P. Dentzien, Trad.). Zahar.

Benjamin, W. (1975). A obra de arte na época de suas técnicas de reprodução. In W. Benjamin, T. W. Adorno, J. Habermas & M. Horkheimer (Orgs.), *Textos escolhidos* (pp. 9-34). Victor Civita.

Berman, M. (1983). *All that is solid melts into air: The experience of modernity*. Verso.

Bengio, Y., Courville, A., & Goodfellow, I. (2016). *Deep learning*. MIT Press.

Bengio, Y., Courville, A., Goodfellow, I. J., Mirza, M., Pouget-Abadie, J., Ozair, S., Warde-Farley, D., & Xu, B. (2014). Generative adversarial network. *arXiv*. https://arxiv.org/abs/1406.2661

Bruner, J. (1956). Freud and the image of man. *Partisan Review*, *23*(3), 340-347.

Chiriatti, M., & Floridi, L. (2020). GPT-3: Its nature, scope, limits, and consequences. *Journal for Artificial Intelligence*, *30*(4), 681-694. https://link.springer.com/article/10.1007/s11023-020-09548-1

Davenport, T. H., & Mittal, N. (2022, 14 de novembro). How generative AI is changing creative work. *Harvard Business Review*. https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work

Flusser, V. (1985). *Filosofia da caixa preta: Ensaios para uma futura filosofia da fotografia*. Hucitec.

Freud, S. (1996a). Uma dificuldade no caminho da psicanálise. In S. Freud. *Edição standard brasileira das obras psicológicas completas de Sigmund Freud: Vol. 17. Uma neurose infantil e outros trabalhos (1917-1918)*. Imago.

Freud, S. (1996b). Introdução ao narcisismo. In S. Freud. *Edição standard brasileira das obras psicológicas completas de Sigmund Freud: Vol. 15. Conferências introdutórias sobre psicanálise (partes I e II) (1915-1916)* Imago.

Freud, S. (1996c). Além do princípio de prazer. In S. Freud. *Edição standard brasileira das obras psicológicas completas de Sigmund Freud: Vol. 18. Além do princípio de prazer, psicologia de grupo e outros trabalhos (1920-1922)*. Imago.

Gagliardi, C. (2022, 16 de dezembro). Os protestos contra o uso da inteligência artificial na arte. *Nexo*. https://www.nexojornal.com.br/expresso/2022/12/16/Os-protestos-contra-o-uso-de-intelig%C3%AAncia-artificial-na-arte

Garrido-Merchán, E. C., & Gozalo-Brizuela, R. (2023). ChatGPT is not all you need: A state of the art review of large generative AI models. *arXiv.org*. . https://doi.org/10.48550/arXiv.2301.04655

Griffith, E., & Metz, C. (2023, 7 de janeiro). A new area of A.I. booms, even amid the tech gloom. *New York Times*. https://www.nytimes.com/2023/01/07/technology/generative-ai-chatgpt-investments.html

Harari, Y., Harris, T., Raskin, A. (2023, 28 de março). Humanidade precisa dominar IA antes que ela nos domine. *Folha de S.Paulo*. https://www1.folha.uol.com.br/tec/2023/03/humanidade-precisa-dominar-ia-antes-que-ela-nos-domine.shtml

Huttenlocher, D., Kissinger, H., & Schmidt, E.. (2023, 24 de fevereiro). ChatGPT heralds an intellectual revolution. *Wall Street Journal*. https://archive.is/2023.02.25-062924/https://www.wsj.com/amp/articles/chatgpt-heralds-an-intellectual-revolution-enlightenment-artificial-intelligence-homo-technicus-technology-cognition-morality-philosophy-774331c6#selection-133.5-631.61

Kaufman, D. (2022a, 19 de agosto). A linguagem natural dos humanos e a linguagem matemática das máquinas. *Época Negócios*. https://epocanegocios.globo.com/colunas/IAgora/noticia/2022/08/linguagem-natural-dos-humanos-e-linguagem-matematica-das-maquinas.html

Kaufman, D. (2022b). *Desmistificando a inteligência artificial*. Autêntica.

Kaufman, D. (2023, 6 de janeiro). As limitações da sensação tecnológica de 2023: o ChatGPT. *Época Negócios*. https://epocanegocios.globo.com/colunas/iagora/coluna/2023/01/as-limitacoes-da-sensacao-tecnologica-de-2023-o-chatgpt.ghtml

Klein, E. (2023, 26 de fevereiro). The imminent danger of A.I. is one we're not talking about. *New York Times*. https://www.nytimes.com/2023/02/26/opinion/microsoft-bing-sydney-artificial-intelligence.html

Kouw, M., Schomaker, L., & Zant, T. (2013). Generative Artificial Intelligence. In Müller, V. (Series Ed.). *Studies in applied philosophy, epistemology and rational ethics: Vol. 5. Philosophy and theory of artificial intelligence* (pp. 107-120). https://doi.org/10.1007/978-3-642-31674-6_8

Mazlish, B. (1993). *The fourth discontinuity: The co-evolution of humans and machines*. Yale University Press.

McLuhan, M. (1969). *Os meios de comunicação como extensões do homem* (D. Pignatari, Trad.), Cultrix.

Butterick, M. (2023, 14 de janeiro). Class action filed against Stability AI, Midjourney, and DeviantArt for DMCA violations, right of publicity violations, unlawful competition, breach of TOS. *Business Insider*. https://markets.businessinsider.com/news/stocks/class-action-filed-against-stability-ai-midjourney-and-deviantart-for-dmca-violations-right-of-publicity-violations-unlawful-competition-breach-of-tos-1032018572

Santaella, L. (2021). Percepção: fenomenologia, ecologia, semiótica. Cengage Learning.

Sloterdijk, P. (2006). Esferas III: Espumas – Esferología plural (I. Reguera, Trad.). Ediciones Siruela.

Uszkoreit, J. (2017, 31 de agosto). Transformer: A novel neural network architecture for language understanding. *Google*. https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html

Vincent, J. (2022, 1º de setembro). An AI-generated artwork's state fair victory fuels arguments over 'what art is'. *The Verge*. https://www.theverge.com/2022/9/1/23332684/ai-generated-artwork-wins-state-fair-competition-colorado

Weber, M. (1967). *A ética protestante e o espírito do capitalismo* (M. I. Q. F. Szmrecsányi & T. J. M. K. Szmrecsányi, Trads.). Pioneira.

Wolfran, S. (2023, 14 de fevereiro). What Is ChatGPT Doing…and Why Does It Work? *Stephen Wolfram Writings*. https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/

Article received on April 20, 2023 and approved on March 5, 2024.