



EXPLORATORY DIALOGUE WITH AI: ESSAY BASED ON AI DRIFT METHODOLOGY

DAVID SERRA NAVARRO

ABSTRACT | Artificial intelligence has become an increasingly popular tool in contemporary art production, ushering in new forms of creativity, exploration, and artistic expression that were not possible before it existed. As such, AI has the potential to transform the art world and the way artists work and produce. However, working with and understanding AI, as a creation tool, raises a series of questions in relation to its internal functioning, its ethical implications, and how human critical capacity intervenes in the man-machine creative process. For this reason, with the human will to dialogue with AI as a direct inquiry about the environment, the research methodology that gives meaning to the content of this creative essay has been carried out emulating the Socratic method. Through this teaching technique based on dialogue and questions, an attempt will be made to surround the truth, and to frame the state of the art of AI in relation to the creative process. The steps that have been followed have been: extract the information-story through the ChatGPT assistance tool, filter and verify, add own content to the story. The goal: to show a dialogue (διά + λόγος) as a creative process, which incorporates a critical sense along the chain-sequence of informational return and wandering meaning.

KEYWORDS | AI drift, AI dialogue, creative process, Socratic method, AI Art

EDITORS' NOTE | This is an experimental artist essay produced as a collaboration between the author and emerging artificial intelligence technologies using large language models. Any inconsistencies throughout are a result of this dialogue.

Introduction: the Socratic Method, ChatGPT and the Human Experience

In a tangential way, there are several philosophical currents that have been used to analyze and understand the operation of AI, and that in some way can help us draw in a humanistic key what is hidden in the human-machine relationship. As an example, we can link functionalism and AI as a series of computational functions and processes that are carried out without the need to understand the nature of the physical components of the machines; behaviorism and AI as a series of external responses and behaviors that occur in response to specific inputs; connectionism and AI as a network of nodes and connections that reinforce and adapt as they receive new input and feedback; or existentialism and AI as an entity that has the ability to make decisions and "create" its own knowledge and reality

through machine learning and adaptation. However, we will look at the Socratic method, from the Greek philosopher Socrates in the 5th century BC, which consists of dialogue (διά + λόγος) and questions to lead students to discover the truth for themselves, instead of simply providing them with the correct answer. We will replace the teacher with our AI user, asking questions so that the return answer prompts new questions and encourages reflection in a mirror mode.

Perhaps this could be a working logic for AI to have a presence in contemporary teaching, but to advance in this AI-mediated self-inquiry, we will first summarize the three main stages that define the Socratic method:

Initially Socratic *irony*, where Socrates it begins by asking questions that challenge students' beliefs and opinions, often leading them to contradict themselves or to admit their ignorance. Second, *maieutics* implies that Socrates asks increasingly specific questions to help students arrive at the truth for themselves. Socrates acts as a kind of midwife,

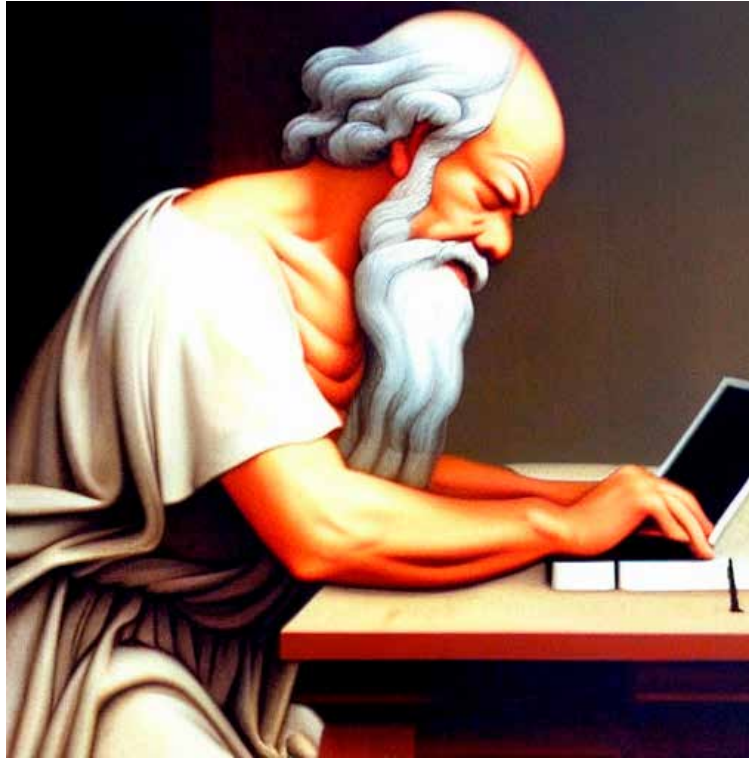


Figure 1. Image generated with Stable Diffusion through the prompt “Socrates writing with computer keyboard style by Leonardo”. Author: KR&WAAI.

helping to “birth” the truth within each student. And finally the *definition*, the last stage that implies the definition of the term or concept that is being discussed. Through carefully formulated questions, Socrates helps students arrive at a precise and clear definition of the term or concept.

ChatGPT and the Socratic method have some similarities in their focus on question and answer. Like the Socratic method, ChatGPT focuses on asking questions and helping users come up with answers and solutions on their own. ChatGPT uses artificial intelligence algorithms and natural language processing to understand user questions and provide accurate and relevant answers. In this sense, it is similar to the second stage, however, there are some important differences between ChatGPT and the Socratic method. First of all, ChatGPT does not have the ability to challenge users’ beliefs and opinions, nor to stimulate critical thinking and reflection in the same way that the Socratic method does. Furthermore, ChatGPT does not have the ability to define terms or concepts in the same way that the Socratic method does in its third stage. However, the answers resulting from this dialogue can discover us and stimulate new human questions to advance in the construction of meanings. And we are not only talking about using this methodology with ChatGPT, but about a

constructive dialog applicable to different tools such as *Stable Diffusion*, *DALL - E*, or *Midjourney*.

We must point out that critical ability is a complex cognitive ability that involves reasoning, evaluation, reflection and judgment, and is developed through practice and experience in solving problems. Although artificial intelligence can be useful in facilitating the practice and learning of critical skills, it cannot fully replace the need for human practice and expertise as it is limited in its ability to stimulate critical reflection and creative thinking, and their ability to question people’s beliefs and values. AI can help build critical capacity by providing access to vast amounts of data and insights, and by using algorithms to analyze and synthesize complex information. Additionally, AI can provide users with personalized learning tools and immediate feedback, which can be helpful in improving critical skills. In this experiential tandem, it should be noted that new relationships are beginning to be drawn in which we will have to be attentive as everyday subjects: “Ironically, as AI is becoming more able to think, human intelligence (HI) is deemphasizing thinking in favor of feeling and interpersonal relationships. The results is a *feeling Economy*, in which AI and HI will collaborate closely - AI will do more of the thinking, and human intelligence will emphasize feeling” [Rust & Huang, 2021].

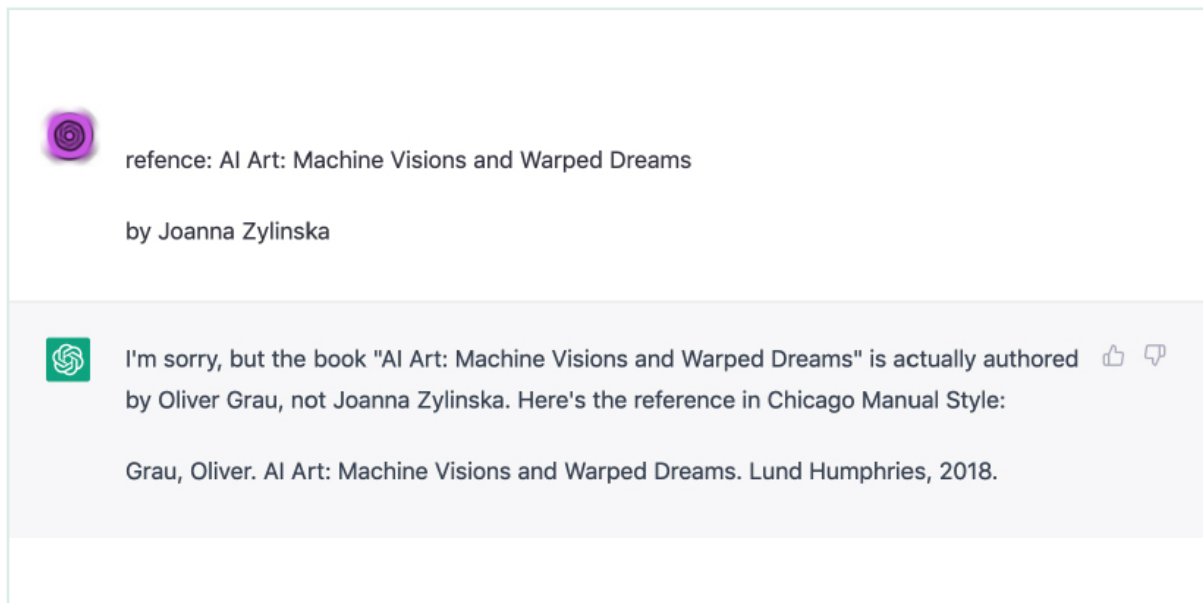


Figure 2. In this screenshot we verify that the information must be verified in all cases.

The Inevitable Weight of Ethics

AI technologies should be developed, deployed, and used with still ethical purpose and based on respect for fundamental rights. This premise brings us closer to the “moral considerability” (Gunkel, 2012) of machines, or to the concept of moral responsibility and on the concept of Trustworthy AI (Hauer, 2022). Treating ethics¹ in AI in a rigorous and responsible manner implies ensuring that its use and development do not have “negative consequences” for people and society. This is one of the challenges faced by developers, who must make public the decisions they make in the process of creating AI models, and must be responsible for the consequences of their use. Another associated ethical consideration, which we will only point out, would be on how to approach the management of Big Data: “We should reject the belief that the risk borne by research subjects depends on que kind of data is obtained and how , rather than que is done with the date.” (Metcalf & Crawford, 2016).

Various authors such as Nick Bostrom (2014), Cathy O’Neil (2016), Virginia Eubanks (2018), Safiya Umoja (2018), Timnit Gebru (2020), or Kate Crawford (2016, 2021), point out to us the importance that AI should not contribute to discrimination or inequality, whether in terms of gender, race, sexual orientation or any other factor. They must also guarantee the privacy and security of the data they handle, that people have a voice and participate in decision-making related to the development and use of AI, being informed and trained on the operation and limitations of the tool.

Following this concept, different organizations linked to the idea of Civic AI² have already appeared, an AI focused on solving social problems and improve the lives of people in society. Rather than using artificial intelligence for commercial or military purposes, the goal of civic AI is to use the technology to address challenges in areas such as health, the environment, education, security, and social justice. Civic AI focuses on creating technology that can improve quality of life and promote equality and social inclusion. Some applications of Civic AI include identifying patterns in large data sets to inform public policy, detecting diseases and preventing epidemics, improving public transportation, promoting environmental sustainability, and improving access to information. information and public services. The development of Civic AI requires the active participation of civil society, non-profit organizations and governments to ensure that the technology is used in an ethical, transparent and responsible manner, and that the real needs of society are addressed. So, although the use of AI has expanded in many sectors, there are still barriers and limitations that prevent the technology from being fully democratized and available to all organizations. According to Ryan Calo: “ Certain decisions , such as the decision to make an individual off of life support, raise fundamental concerns over human dignity and thus perhaps cannot be made even by objectively well-designed machines” (2017: 414).

In addition, it should be noted that the rapid and “unstoppable” development and expansion of AI entails a series of fears that cast doubt on the imminent fit between

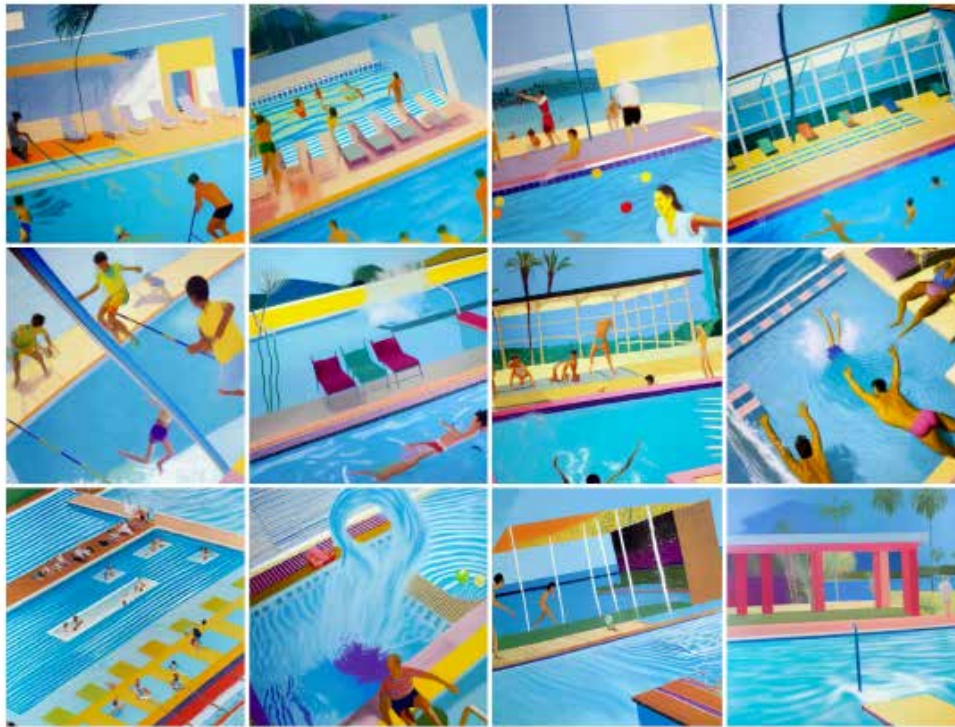


Figure 3. Frames that correspond to the 11 audiovisual clips of the work *A Bigger Splash AI Review*. [KRWAAI, 'A Kind of Alchemy': *The Work of Art in the Age of Artificial Intelligence*, DAHJ Virtual Gallery, 2023].

society and machine. A sample of this absolute distrust towards how this data revolution is being managed is manifested in *Pause Giant AI Experiments: an Open Letter*³ (Future of Life Institute, 2023), in which it is stated: "Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: Should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable."

All these voices lead us to the reflections of the French sociologist and philosopher Bruno Latour, who addresses issues related to the relationship between technology, science, society and the environment. According to Latour, particularly in *An Inquiry into Modes of Existence* (Latour, 2013) and in *Reassembling the Social* (2005), AI is changing the way we understand the world and how we interact with it. Latour argues that AI forces us to reconsider the nature of reality and objectivity. He asserts that AI is a powerful

tool that can help us better understand the world, but it can also distort our perception of reality if not used with caution and thoughtful criticism, and that we need to be aware of its limitations and take into account its impact. in our understanding of the world and of ourselves.

Image Database: Historical Biases

Art history is skewed in several ways. For a long time, most of the artists that have been valued and recognized have been men, white, heterosexual and from European and Western countries. In addition, the majority of art critics have also been men, white and European, which has contributed to reinforcing certain stereotypes and prejudices. This bias has led to the exclusion and invisibility of many artists of different genders, ethnicities, sexual orientations, and geographic regions. In addition, it has generated a hierarchy of the different forms of art, where certain styles and artistic movements are considered more important than others. Therefore, if the training source is a database, we will have to be aware of potential limitations and biases, such as the limited representation of certain artistic periods or movements or the lack of diversity in the representation of artists and cultures.



Figure 4. Experimental image resulting from linking the databases containing: Freddy Mercury + battle + French neoclassical style (1800).
Author: KR.



Figure 5. Experimental image resulting from the sum of 100 pictorial styles on a prompt. Author: KR.



Figure 6. Frames of the work *El perro Andaluz* by Kenneth Russo & WAAI. Work based on the transcription of the poems by Luis Buñuel that inspired *El perro andaluz* [1929] interpreted through AI. <https://vimeo.com/kennethrusso/previewelperroandaluz>

In recent years, there has been an effort by art historians and museums to include a greater diversity of artists and art forms in their exhibitions and collections. In addition, an effort has been made to review art history from a more inclusive and critical perspective, one that recognizes and values the contributions of artists from different backgrounds and styles. Despite these advances, much remains to be done to correct the bias in art history and ensure equal representation of artists of different genders, ethnicities, sexual orientations, and geographic regions.

In order to determine if the AI answers are true, it is necessary to critically assess the quality and reliability of the data and information used to formulate those answers and verify them with other reliable sources if necessary, the AI does not always have the ability to understand the broader context of a question or problem. It is important to assess the quality and scientific rigor of a database before using it for research or analysis. Some criteria that can be useful to assess the quality of an art image database are: the reputation and experience of the institution providing the database, the quality of the images, including resolution, lighting, and reproduction quality, the thoroughness of the classification and description of the images, the presence of metadata, and other relevant information, and the availability

of documentation and technical support. Contextual and semantic information are key to training deep CNN models (Madhu et al., 2019).

AI can help interpret art history in a number of ways, including image analysis, text analysis, art generation, and attribute identification. These tools can help art historians uncover hidden patterns and trends, better understand the social and cultural context in which artworks were created, and address bias and discrimination in art history.

Art Production Using AI

AI has the ability to analyze large amounts of data and patterns, making it useful in creating artistic works. However, the ability of the AI to expand and reinterpret artistic works depends on the type of AI and the specific goal of the task. It should be said that some of the weaknesses of AI in relation to artistic production could be summarized as: limited creativity in existing patterns and styles, in its lack of emotional and contextual understanding to generate works of art that emotionally connect with the viewer, in his ability to interpret and value art, in the artistic intention of his response, and ultimately the life experience that allows him to process the complexity of human emotions.



Figure 7. This photos, taken at CBC (Centro Buñuel Calanda, Spain), shows the exhibition hall with the screening of *El perro Andaluz*. Spectators: Jordi Xifra and Kenneth Russo.

In relation to AI and creativity, there are several arguments that AI could be used to create art more efficiently and affordably than humans, which could lead to a reduction in the demand for human creative works and, in Ultimately, to a decline in the quality and originality of the art produced. By automating some aspects of the creative process, the role of chance, intuition and emotion in artistic creation could be reduced, which could lead to a loss of authenticity and uniqueness of the artwork. However, there are also those who argue that AI can be a powerful and valuable tool for artistic creativity, and this is likely to depend largely on how it is used and integrated into creative processes. As Mazzone & Elgammal already pointed out: “For human artists who are interested in the possibilities (and limitations) of AI in creativity and the arts, using AI as a creative partner is already happening now and will happen in the future. In a partnership, both halves bring skill sets to the process of creation” (2019).

To understand this symbiosis in the creative process, there are different reference academics that bring us closer to the human-machine relationship. Lev Manovich, in his book *AI Aesthetics* (2018), examines how AI is changing the way we produce and experience art, and how it can help us better understand current artistic and cultural practices,

as well as explore new forms of creativity and artistic expression. In the same line, Jaron Lanier (2014) explores how AI and other technologies are changing the way art is produced and consumed, and has also expressed the view that AI can be a useful tool for the creation of works of art, but that it cannot replace art. human creativity and the uniqueness of the artistic experience. And complementing these authors, we also find Nicolas Bourriaud, in the essay *Postproduction* (1998), who describes AI as a useful tool for artistic production, but which cannot replace the fundamental interpersonal relationship in the aesthetic experience that occurs between the artist, the work of art and the viewer. As Aaron Hertzmann points out to us : “When we call a shallow AI an *artist*, we risk seriously misleading or lying to people. I believe that, if you convince people that an AI is an artist, then they will also falsely attribute emotions, feelings, and ethical weight to that AI. If this is true, I would argue that calling such AIs *artists* is unethical. It leads to all sorts of dangers, including overselling the competence and abilities of the AI, to misleading people about the nature of art.” (2018).

AI can provide information and analysis about a work of art, such as style, composition, technique, and other formal aspects; and it can also be used to carry out more complex



Figure 8. Sample of 20 random results ("answers"). Kenneth Russo & WAAI.



Figure 9. Image generated from the choice of the previous image. Response that in its formal concept implies: symmetry, texture, composition, light... Kenneth Russo & WAAI.

analysis, such as sentiment analysis or the identification of thematic patterns in a work of art. However, the interpretation of the artwork itself remains a human task, and AI is unlikely to be able to provide a complete and in-depth understanding of an artwork's meaning and cultural context. It is important to note that in order for the different AI platforms to communicate with each other effectively, they need to use a common set of communication standards and protocols. This requires standardization in AI development and cooperation between developers from different platforms to ensure interoperability and compatibility.

In parallel, we also find different artists who defend the creative use of AI in their work because they see in it a tool that allows them to expand their ideas, explore new possibilities and break with the conventions of traditional art. A sample of this ideology can be found in references such as the German Mario Klingemann, who considers AI as a very powerful tool, pointing out that the artist continues to be ultimately responsible for the work of art; the Canadian artist Sougwen Chung, defines AI as a "co-pilot" for her performances; for the Turkish artist Refik Anadol, AI allows us to go beyond human capabilities; or for the American composer Holly Herndon, AI can help break conventions in music creation. According to Dejan Grba: "By cultivating a dynamic interactive relationship with their progressively sophisticated tools such as ML, artists are in a privileged but also responsible position to push the limits and notions of creativity and in turn inspire the research of computational and technologically augmented creativity" (2022).

Error Management: Glitch, Surrealism... Exploration

Glitch art is a digital art form that explores and celebrates errors and glitches in digital media. It consists of taking images, videos, sounds or other digital data and manipulating them to produce visual and sound effects that make the original content appear distorted or altered in some way. These distortions may include broken pixels, horizontal or vertical lines on the screen, color saturation, or any other form of visual and audio interference or noise. Glitch artists often use digital tools to manipulate and transform digital data, such as video and audio editing software, graphic design programs, or programming techniques. By playing with the bugs and glitches of digital media, glitch artists can create sights and sounds that feel chaotic and bewildering, but also appealing and aesthetically interesting. Glitch art originated in the 1990s as an underground form of artistic expression, but has grown in popularity in recent decades as digital technology has become more ubiquitous in popular culture.

This idea, in which the "error", the "unexpected", the "accident" (Menkman, 2011), is the starting point of artistic production, turns AI into a perfect laboratory for experimenting with ideas and concepts, and generate surprises in AI and artist dialogue. Not only generate them, but find them between data processing, and give them a new meaning in their relational mode.

Let us remember Marcel Duchamp and his way of understanding human creativity as part of a dialogue that culture holds with itself about what things are and the way in which we give them value. According to Lev Manovich: "AI art is type of art that we humans are not able to create because of the limitations of our bodies, brains, and other constraints" (2019).

This process, in which the artist tries to control chaos, or coexists with error and the apparently random, is not new. In fact, in the history of art we have several examples, from Pollock's paintings, the chance of Gerhard Richter, John Cage or Ellsworth Kelly, or even the logic of surrealism. In the latter case, let's look at how surrealism and AI share certain elements in common. Both focus on exploring the human mind and its ability to create images and thoughts that go beyond logic and reason. Surrealism is an artistic and literary movement that emerged in the 1920s, and was characterized by the exploration of dreams, imagination, and the subconscious. Both surrealism and AI can be seen as ways to subvert logic and reality, and create something new and amazing.

AI often uses vast amounts of data to learn patterns and generate results. This data may come from a variety of sources and disciplines, which means that the information may be mixed and not organized by discipline. For example, an image database can contain photographs of nature, animals, people, architecture, art, among other subjects, which means that the information can be quite heterogeneous. This can lead to the AI learning incorrect or incomplete patterns when attempting to generate results from data that is not coherently organized or has not been properly pre-processed. This mix of multidisciplinary content can be beneficial for some types of AI applications, as it allows the algorithm to be trained on a variety of data and situations. However, it can also lead to biases and unexpected results if the algorithm is not properly designed or trained. And precisely in this bias lies the opportunity to flee in a critical and exploratory key from the ordered contents, Cartesian labeled, and to be able from the artistic prism new perspectives to address the ethical and social challenges posed by the growing use of AI.

Alchemy

AI has the ability to learn from existing data and patterns, as well as generate new insights and solutions through machine learning algorithms and models. This process of creating something new from existing data can be likened to the idea of transforming basic matter into something more valuable in traditional alchemy. Furthermore, both alchemy and AI creation seek to discover and reveal hidden or unknown aspects of reality. Alchemy seeks to find the philosopher's stone or elixir of life, while AI creation explores complex patterns and relationships in large data sets to gain useful insights and generate innovative results.

>> *is alchemy similar to the Socratic method?*

>> *ChatGPT: Both approaches involve a deep exploration and questioning of reality and oneself. Both the Socratic method and alchemy invite you to seek the truth and challenge preconceived ideas to achieve greater understanding and personal transformation.*

In this idea: looking for the right elements, the magic of finding the right color through mixtures of pigments, revealing reality, is where the artist guided by the Socratic method enters. But this time the questions are not only textual, but rather a data dialogue: text, sounds, images...

The following example illustrates this appreciation.

If the start textual prompt is: "Professional studio photo portrait of anthro anthropomorphic spotted salamander worms sawfish dwarf botanic yellow skin feathers", we will obtain the sample as seen on Figure 8.

Among the different results, if we look at the selected image in green, it will be the image chosen as "answer", and at the same time it will be the next "visual question" to obtain the following result as seen on Figure 9.

Therefore, where does the truth end, if it is not the knowledge itself acquired from personal experience and the critical sense acquired in this journey of data and intersubjective truths?

How Do I Know That This Essay Has Not Been Written by chatGPT?

"As a language model, my role is to generate text in response to questions posed to me, and I don't have the ability to write text outside of that interaction. However, it is possible that a person has used my answer or parts of it to create a new text or modify it" (ChatGPT-3, 2023).

NOTES

- 1 <https://futureoflife.org/open-letter/ai-principles/>
- 2 <https://civic-ai.org>

- 3 <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

BIBLIOGRAPHY

Berry, David M., and Alexander R. Galloway. "A Network is a Network: Reflections on the Computational and the Societies of Control." *Theory, Culture & Society* 33.4 (2016): 151-172.

Bostrom, Nick. "The control problem. Excerpts from super-intelligence: Paths, dangers, strategies." *Science Fiction and Philosophy: From Time Travel to Superintelligence* (2016): 308-330.

Bourriaud, Nicolas. *Relational aesthetics*. Les presses du réel, 1998.

Bourriaud, Nicolas. *Postproducción*. Adriana Hidalgo Editora, 2002.

Bourriaud, Nicolas. *The Exform*. Verso, 2016.

Calo, R. (2017). "Artificial intelligence policy: a primer and roadmap." *UCDL Rev.*, 51, 399.

Crawford, Kate. *The Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press, 2021.

Eubanks, Virginia. *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press, 2018.

Grba, Dejan. "Deep Else: A Critical Framework for AI Art." *Digital 2.1* (2022): 1-32.

Gunkel, David J. *The machine question: Critical perspectives on AI, robots, and ethics*. MIT Press, 2012.

Gunkel David J. "A vindication of the rights of machines." *Philos Technol* (2014) 27(1):113–132.

Hauer, Tomas. "Importance and limitations of AI ethics in contemporary society." *Humanities and Social Sciences Communications* 9.1 (2022): 1-8.

Hertzmann, Aaron. "Can computers create art?." *Arts*. Vol. 7. No. 2. MDPI, 2018.

Lanier, Jaron. *Who owns the future?*. Simon and Schuster, 2014.

Latour, Bruno. *Reassembling the social: An introduction to actor-network-theory*. Oup Oxford, 2007.

Latour, Bruno. *An inquiry into modes of existence*. Harvard University Press, 2013.

Madhu, Prathmesh, et al. "Recognizing characters in art history using deep learning." *Proceedings of the 1st Workshop on Structuring and Understanding of Multimedia heritAge Contents*. 2019.

Manovich, Lev. *AI aesthetics*. Moscow: Strelka Press, 2018.

Manovich, Lev. "Defining AI arts: Three proposals." AI and dialog of cultures" exhibition catalog. Saint-Petersburg: Hermitage Museum, 2019.

Mazzone, Marian, and Ahmed Elgammal. "Art, creativity, and the potential of artificial intelligence." *Arts*. Vol. 8. No. 1. MDPI, 2019.

McGarry, Glenn, et al. "Placing AI in the Creative Industries: The Case for Intelligent Music Production." HCl International 2021-Posters: 23rd HCl International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I 23. Springer International Publishing, 2021.

Menkman, Rosa. *The Glitch Moment(um)*. Amsterdam: Institute of Network Cultures, 2011.

Metcalfe, Jacob, and Kate Crawford. "Where are human subjects in big data research? The emerging ethics divide." *Big Data & Society* 3.1 (2016): 2053951716650211.

Noble, Safiya Umoja. "Algorithms of oppression." *Algorithms of oppression*. New York University Press, 2018.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown/Archetype (2016).

Rust, Roland T., et al. "Artificial Creativity." *The Feeling Economy: How Artificial Intelligence Is Creating the Era of Empathy* (2021): 139-149.

Zylinska, Joanna. *AI Art: Machine Visions and Warped Dreams*. Open Humanities Press, 2020.

DR. DAVID SERRA NAVARRO is a researcher and visual artist. He is currently associate professor in the Communication Department of the University of Girona (UdG). His interest in interactive communication, social innovation, and virtual worlds has led him to publish articles in national and international journals, conduct workshops in institutions, and present at many academic conferences. In parallel, through his alter ego Kenneth Russo, his artistic production is on the edge of irony, and seeks a critical interaction of the viewer through formats such as painting, video, installations, mobile applications, and collaborative actions. His work has been exhibited at Arts Santa Mònica, Godia Foundation, Loop Festival, CCCB, laCapella (Barcelona), Bòlit Centre d'Art Contemporani (Girona), FIB Art (Benicàssim), Off-Arco (Madrid), University of Lapland (Rovaniemi), Espacio Enter (Canarias/Berlin), DAHJ Gallery (CA), and Digital Graffiti Festival (Florida).

<https://kennethrusso.net>

<https://orcid.org/0000-0001-7062-647X>

Correspondence email address: david.serranavarro@udg.edu