Generating a living creature?

Plato on writing and two kinds of generation

HUA-KUEI HO
Chinese Culture University

Abstract

Generative AI can generate texts, images, and other kinds of content efficiently. When you ask any question, it can generate various answers, just like a living thing. In a sense, it sounds like Plato's ideal writing. In the *Phaedrus*, Plato criticises writing and compares it to an offspring that "always needs its father's help" (275d4-e5). Writing presents information as if it were considering something with intelligence. However, it lacks genuine life. One sign of writing's lifelessness is its inability to respond to questions. In contrast, Plato's ideal form of writing is like generating a living creature (ζῷον, 264c3). Not every generation gives birth to fertile children. Plato discriminates between two kinds of generation and connects the metaphor to the issue of discerning truth from falsehood in the *Theaetetus*. In this paper, I will explore this connection. I will argue that Plato does not mind the source of information. Still, he is really concerned with the process of examining truth in any given data. To examine what is generated, we need a partner to discuss together, exchange questions and answers, and, most importantly, make ourselves engaged. Why do we prefer human partners in conversations? Writing, generative AI, and other information technologies may also play the role of our partner. Nonetheless, if we do not urge ourselves to be engaged in the questioning and answering, we are affected by information passively, and we are vulnerable to being manipulated by human-generated writing or AI-generated content. Plato's critique of writing is not a rejection of new technology. Rather, the philosopher uses writing, the newly developed information technology of his time, to equip us with the ability to resist information manipulation. This ability is essential if we are to have real communication and to approach knowledge.

Keywords: Plato, writing, generative AI, *Phaedrus*, *Theaetetus*

Generating a living creature?

Plato on writing and two kinds of generation

1 Revisiting Plato in the age of generative AI

Plato's criticisms of writing in the *Phaedrus* are well-known.[1] He criticises writing for inhibiting memory; written words are fixed and not open to clarification; furthermore, written words are unable to adapt to different audiences. He compares the relationship between writing and its author to that between offspring and their parent. As Plato wrote, Socrates said the "offspring" of writing would be exposed to misunderstandings. When you ask it any question, it "always refers to the only single same answer" and "always needs its father's help" (*Phaedrus*, 275d4-e5).[2] Despite this, Plato wrote a lot. He must have some reason to convince himself to do so.[3] This leads his readers to wonder what kind of writing he would accept.[4] Earlier in the dialogue, Plato describes ideal writing as composing something like a living creature (ζῷον) with its organic structure (264c2-5). Ideally, the writing process resembles the generation of a living being.

Generative AI enables its users to generate new texts, images, and other kinds of content. When you ask it any question, it can generate various answers, just as a living thing, without depending on its "father". Does this sound like that generative AI makes Plato's ideal writing come true?

---

2   Greek citations in this paper are according to Oxford Classical Texts. Translations are mine unless otherwise noted.

3   Since the *Phaedrus* itself is written, it is natural to apply the critique to Plato's own writing. For the self-reference of Plato's critique, see Hackforth, 1952, p. 164; Rowe, 2007, p. 271; and Long, 2013, p. 16, among others.

4   This is not a new topic. A typical response is that Plato uses writing or any medium "that typically mimics the conditions of oral communication, so far as it can" (Rowe, 2015, n. 39; for more details, see Rowe, 2007). However, I will not endorse the superiority of the oral medium in this paper. The focus will turn to the generative aspect of writing.

The metaphor of generation is not unfamiliar to Plato's readers. In the so-called Socrates' midwifery, the philosopher attempted to help "pregnant" young people give birth to knowledge (*Theaetetus*, 148e7-151d6). However, what is generated must be examined, whether it is fertile or a wind-egg (151e5-6). There seem to be two kinds of generation. Plato adopts one but rejects the other. Is it fertile or a wind-egg? It is the question.

In this paper, I will revisit Plato's critique of writing in the *Phaedrus*, in the light of the two kinds of generation implied in the midwifery story in the *Theaetetus*, to clarify what Plato aims to generate by writing. Surely, there are some problematic issues he worries about. Nevertheless, Plato's critique of writing can be taken as a fruitful reflection on writing, not an indiscriminating rejection. Employing the newly developed information technology, albeit with great caution, the philosopher strives for better communication, which hopefully will achieve real knowledge rather than simply spreading information.

2 Generating a living creature by writing

In the *Phaedrus*, when Socrates criticises the disorder of Lysias' written speech, which Phaedrus exhibited to him, he submits the idea that writing should be like a living creature:

> But, at least, I think you would say that every speech (λόγον) ought to (δεῖν) be composed (συνεστάναι) as a living creature (ζῷον), which has a body of its own; it needs to be neither without a head nor without feet, but have both middle parts and extremities, fitting to each other and to the whole when they are written.
>
> (Plato, *Phaedrus*, 264c2-5)

It is a normative claim. Every speech ought to be composed as a living creature—ζῷον in Greek, a word derived from the verb ζάω (to live), means something with life by definition. So, the head, feet, middle

parts, and extremities do not refer to disparate portions of a non-organism, nor limbs of a corpse. It needs to be a living whole. Although, in the context, the target of criticism is the disorder of Lysias' speech, it goes beyond the question of arranging things in an orderly way. It is a question of generating something alive.

When generated, will this "living creature" be immune to Plato's critique of writing? Socrates in the dialogue explains the weaknesses of writing[5] as follows:

> It's because, O Phaedrus, writing has this smart feature that is truly the same as painting (ζωγραφίᾳ). The offspring of the painting stand as if they are alive (ζῶντα), but if you ask any question, they totally keep solemnly silent. Words are the same. You think they are speaking (λέγειν) as if they are considering something with intelligence (τι φρονοῦντας), but if you ask any of what is said when you want to learn (μαθεῖν), it always refers to the only one same answer. Once it has been written down, every word rolls everywhere all the same, to those who listen (τοῖς ἐπαΐουσιν), in the same way to those who have nothing to do with it, and it doesn't know (οὐκ ἐπίσταται) to whom it should speak and to whom it should not. When it is misunderstood and abused not justly, it always needs its father's help; it cannot defend itself nor help itself.
>
> (Plato, *Phaedrus*, 275d4-e5)

Writing is like painting. Their offspring "stand as if they are alive", but they are not really with life. What are the signs showing them lifeless? They cannot really answer questions, and they merely repeat the same words. They cannot adapt to different readers but are open to misunderstandings. They depend on their parent, the author, when confronted with misinterpretations.

---

5   I skip at least one important weakness: writing makes people forgetful, as shown by the story of Theuth's invention of writing (*Phaedrus*, 274c5-275b2). The weakness is connected to the comparison of writing to a *pharmakon*, which once attracted Derrida's attention (Derrida, 1972). The use of *pharmakon* is rich in controversy in Plato. It is too complex to be discussed here.

In the citation, Plato also uses λέγειν (to speak)—a term which covers both orality and writing, and ἐπαΐουσιν (listening)—a term which is usually associated with orality. There is no absolute distinction between written and spoken words.[6] As Trotz-Liboff argued in his article on Plato's comparison of writing to a *pharmakon*, "speech as spoken, rather than written, does not automatically overcome the problems of communication: oral speech still requires interpretation and can demonstrate the defects of writing" (Trotz-Liboff, 2023, p. 400). More broadly, Plato's reflection here can be applied to painting, writing, speaking, and all kinds of information technology, all media of communication.

Talking about answering questions and not repeating the same words, this is what generative AI can do very well. Generative AI is "artificial intelligence that is capable of generating new content (such as images or text) in response to a submitted prompt (such as a query)" (Merriam-Webster, 2024). Suggesting you to change words has been a very basic function for AI writing tools. As for responding to different people in varied ways, generative AI can interact with users to a certain degree now, and aims to do it better and better.[7] Given sufficient time and data, it is highly probable that it will provide more individualised answers. This is because generative AI does not learn a fixed set of answers. Instead, it mimics the human learning process.[8] It can learn to adapt its responses to individual users.

Technical writer Margaret Rouse provides a helpful comparison of artificial intelligence (AI), traditional machine learning (ML) and generative AI (Rouse, 2024):

---

6   For Plato's transgression of the boundary between speaking and writing, see Ho, 2011.

7   The leading technology company in the AI boom Nvidia marks out "three key requirements of a successful generative AI model": Quality, Diversity, and Speed. The requirements arise from the uses of "interactive applications" (Nvidia, 2024). The interaction with users is taken into consideration crucially.

8   "One of the breakthroughs with generative AI models is the ability to leverage different learning approaches, including unsupervised or semi-supervised learning for training" (Nvidia, 2024).
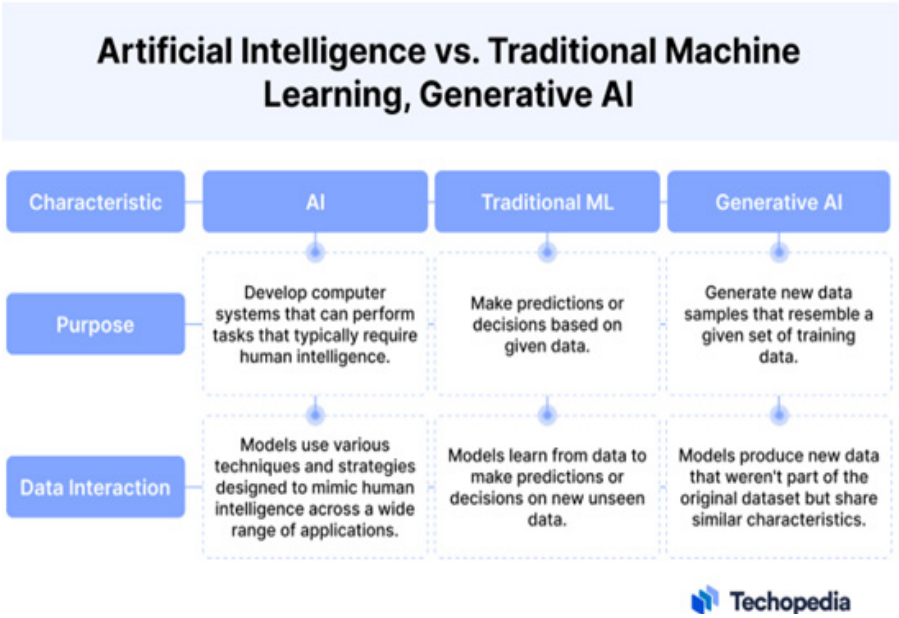
Artificial Intelligence vs. Traditional Machine Learning, Generative AI

Characteristic	AI	Traditional ML	Generative AI
Purpose	Develop computer systems that can perform tasks that typically require human intelligence.	Make predictions or decisions based on given data.	Generate new data samples that resemble a given set of training data.
Data Interaction	Models use various techniques and strategies designed to mimic human intelligence across a wide range of applications.	Models learn from data to make predictions or decisions on new unseen data.	Models produce new data that weren't part of the original dataset but share similar characteristics.

Techopedia

Table 1: Comparison of AI, ML and Generative AI

It shows that the direction of the development of AI has turned imitative learning into a more interactive generating.

However, the problem of misunderstanding remains. The problem is shared by writing, outputs of AI, and also among people. Even if a newly developed information technology can mimic human intelligence perfectly, responding to different individuals differently, the problem of misunderstanding remains.

When Plato says written words need their father's help, does he believe the "father" can really defend against misunderstandings? The author who generates writing, or whatever which generates any information, is not as significant for Plato as expected. Authorship, if understood as the source of written information, is not Plato's concern. Just a moment before comparing writing to painting, Phaedrus alertly noticed that Socrates had made up the story of the invention of writing. Socrates did not deny this, but he argued that ancient people "listen to an oak or

a rock due to their good nature, *as long as it tells the truth*; but perhaps it makes a difference for you, who is the person speaking and where he comes from" (*Phaedrus*, 275b8-c1). He does not care about the source of information "as long as it tells the truth". His concern is truth in the speech's content. Nevertheless, if truth does not reside in the source, how shall we distinguish truth from falsehood when we are fed by information?

3 Two kinds of generation

Plato extends the comparison of writing to offspring beyond the *Phaedrus*. This metaphor also features in the *Theaetetus*. Specifically, Protagoras' work, *Truth*, is depicted as "an orphan". This portrayal stems from the fact that Protagoras, its author, is dead and, therefore, unable to defend his written words (*Theaetetus* 164e2-7). We will now shift our focus to the *Theaetetus*, a dialogue that delves into the nature of knowledge. It articulates the connection between the metaphor of generation and the issue of discerning truth from falsehood.

In the renowned passage regarding the midwifery of knowledge in the *Theaetetus*, Socrates asserts that his work surpasses that of midwives. Women don't sometimes "generate shadows" (εἴδωλα τίκτειν), sometimes "true children" (ἀληθινά). It is not easy to distinguish (διαγνῶναι) between these (150a8-b4). Unlike women's midwifery, the most important task for Socrates' profession (τέχνη) lies precisely in discerning genuine offspring from mere illusion. He needs to test by every method whether the young man's thought is generating (ἀποτίκτει) "a shadow and falsehood" (εἴδωλον καὶ ψεῦδος) or "a fertile and true child" (γόνιμόν τε καὶ ἀληθές, 150b9-c3).

This highlights two kinds of intellectual generation: one yielding a true, fertile child, representing truth in thought, and the other producing a shadow, a false offspring, representing falsehood. Shortly after this assertion, when Theaetetus brings forth his thought, Socrates says, "Let's examine it together" (αὐτὸ κοινῇ σκεψώμεθα) whether it is "fertile or a

wind-egg" (γόνιμον ἢ ἀνεμιαῖον), to practice his midwifery on Theaete-tus (151e5-6). If the thought cannot pass the test, it is "only a wind-egg and false" (ἀλλὰ ἀνεμιαῖόν τε καὶ ψεῦδος, 161a1-4). It will be taken away.

In the rest of the dialogue, we see a long process of questioning and answering between Socrates and Theaetetus. Such a question-and-an-swer exchange is an indispensable feature of philosophy. Burnyeat read Socratic midwifery as "an account of a method of education which is at the same time a method of doing philosophy". As he pointed out, "its procedure should be a discussion in which Socratic questioning engag-es with one's own personal conception of things" (Burnyeat, 1990, pp. 6-7). The continuous question-and-answer exchange between Socrates and Theaetetus illustrates how they decide which kind of generation it is whenever an answer is generated. If what is generated is false, it is a "wind-egg", and they will abandon it.

This is consistent with the requirement of truth, as suggested at the end of the previous section. Socrates' midwifery does not care "who is the person speaking and where he comes from", namely, the parent or source of information. Plato can accept words coming from an oak or a stone "as long as it tells the truth" (*Phaedrus*, 275b8-c1). Accordingly, words coming from a generative AI can be reasonably accepted, on the condition that *they tell the truth*. This echoes Plato's attitude of not car-ing about bloodline, as shown in the "digression" passage of *Theaetetus*. He devalues the praises of lineage or noble birth as narrow-visioned. The praises come from those who cannot see that everyone has count-less "grandfathers and ancestors", including rich people, poor people, kings, slaves, barbarians and Greeks (*Theaetetus*, 174e5-175a5). The point is whether *it tells the truth*. The parents, grandparents and ances-tors of what is generated can be an abundant mixture.[9] It does not mat-ter. The truth does not rely on the authority of the source of information, but calls for the examination consisting of questioning and answering.

---

9   "Generative AI enables users to quickly generate new content based on a variety of inputs" (Nvidia, 2024). "[T]his type of AI learns patterns from training data and generates new, unique outputs with the same statistical properties" (Rouse, 2024). The source-data must be a huge mixture.

The practice of questioning and answering in midwifery parallels Plato's critique of writing in the *Phaedrus*. One key sign of writing's lifelessness, as noted in the *Phaedrus*, is its inability to respond to questions. Written words present information "as if they are considering something with intelligence" (275d7), yet they cannot engage in real dialogue. Given this context, the ability to answer questions is a necessary, but not sufficient, condition for Plato's ideal writing. While generative artificial intelligence excels at this particular task, it does not fully meet Plato's criteria. If the significance of Socrates' midwifery lies in discerning truth from falsehood through a process of questioning and answering, what, in this process, truly reveals the truth?

4 Truth, knowledge, and communication

The meaning of "telling the truth" is not intuitive. Truth in Plato is complex. Socrates makes up the story of the invention of writing in the *Phaedrus*. A more notorious case is the use of "noble lie" in the *Republic* (414b8-415c7). In Plato's own words, the made-up stories or lies are "as a whole telling false, but there is something true in them" (*Republic*, 377a5-6). How do we recognise this "something true"? Suppose an oak or a stone tells us something true (cf. *Phaedrus*, 275b8-c1). According to the midwifery of knowledge, we cannot accept any proposition as true without examination, regardless of its apparent truth. Truth cannot be the sign of life. The offspring of writing, or any medium of information, cannot be accepted as true and fertile unless we have examined it. To decide which kind of generation it is, we need to look for the sign of life in the process of questioning and answering.

This is not to ask for countless questions and answers, but to do philosophical discussion. In the discussion, "questioning *engages with* one's own personal conception of things" (Burnyeat, 1990, p. 7, italics mine). Plato's writing takes the form of dialogue, and his philosophy is widely recognised as "dialogical". Long clarifies that such a philosophical activity can be an interpersonal conversation between people or an internal dialogue undertaken solo (Long, 2013, pp. 3-6). He observes

that there is "a pervasive theme in Plato's accounts of dialectical in-quiry: the need for a suitable partner" (p. 24). Studying the *Theaetetus* and *Sophist*, he argues that even when the author of written words is physically absent, it is still possible to do self-critical thinking through internal dialogue (pp. 126-138). For instance, the imagined conversation with Protagoras in the *Theaetetus* demonstrates the possibility. To examine words in the *Truth*, Protagoras' offspring of writing, Socrates imitates Protagoras to respond to his questioning (*Theaetetus*, 170a6). He further imagines that Protagoras popped up from the ground to his neck as if he became alive again to deny Socrates' words (171d1-2; cf. Long, 2013, pp. 136-137).

In short, the examination of what is generated needs to be done "together" (κοινῇ, *Theaetetus*, 151e5). We need a suitable partner to discuss together, exchange questions and answers together, and *make ourselves engaged*.

Nowadays, one may undertake a philosophical discussion with a very different partner. In a recent paper on ChatGPT-3 and Plato's critique of writing in the *Phaedrus*, Loos and Radicke recorded their question-and-answer exchange with generative AI. We see that generative AI can do self-reflection and even philosophical reflection (Loos & Radicke, 2024, pp. 9-11). An attention-grabbing experiment they conducted was to ask ChatGPT-3 to reflect on the similarities between Plato's critique of writing and ChatGPT-3 as a writing tool.[10] ChatGPT-3 catches three important concerns in Plato: the impact of writing on human memory, knowledge acquisition, and the development of critical thinking skills, and confirms that these concerns are also related to itself (pp. 10-11). Writing, generative AI and other information technologies may play the role of our partner in conversations. Nevertheless, seeing your partner exhibiting questions and answers is not the same thing as being engaged in the process. If we are not engaged in the questioning and answering, we lose the opportunity to give birth to true fertile offspring, and the whole process will have nothing to do with our own knowledge.

---

10   If one prefers a similar reflection provided by a human being, Roochnik argues that Plato conceives of writing as a kind of technology. The critique of writing in the *Phaedurs* can be applied to our most recent technologies, including AI (Roochnik, 2024).

Besides, this "partner" can be manipulated by a powerful company or even by a government. Many people are worrying that generative AI can be an efficient tool to manipulate us. Toxic manipulation is not an innovation. If we do not urge ourselves to be engaged in the questioning and answering but only receive information passively, we are vulnerable to being manipulated by human-generated writing or AI-generated content.

Interestingly, Loos and Radicke note that ChatGPT-3 does not consider some points that are important to Plato. One among them is the problem of misunderstanding (p.11). I suggest that ChatGPT-3 is smart and correct in ignoring the problem of misunderstanding. As argued in Section 2, the problem of misunderstanding is shared by writing, AI, and human beings. Words from a living human being do not "automatically overcome the problems of communication" (Trotz-Liboff, 2023, p. 400, cited in Section 2). Indeed, no matter how capable of answering questions a generative AI is, what is generated cannot prevent being misunderstood. However, the problem should not be attributed to writing, generative AI writing tools, or any information technology. Not only can dead words be misunderstood, but a living human being can also be misunderstood. Misunderstanding is an enduring problem for communication.

The only way to face the problem is to continue to communicate. As a philosopher, Plato takes responsibility for facing the problem through a remarkable number of dialogues. Despite the possibility of being misunderstood, he employs writing, the information technology of his days, to converse with himself as well as us, cautiously but engagingly. This kind of generation is definitely fertile.

References

Burnyeat, M. (1990). *The Theaetetus of Plato*, with a translation by M. J. Levett, revised by M. Burnyeat. Hackett.

Derrida, J. (1972). La Pharmacie de Platon. *La Dissémination* (pp. 69-199). Seuil; B. Johnson (Trans.) (1981). Plato's Pharmacy. In *Dissemination* (pp. 61-171). University of Chicago Press. https://doi.org/10.7208/chicago/9780226816340.001.0001

Hackforth, R. (1952). *Plato's Phaedrus*, translated with an introduction and commentary. Cambridge University Press.

Ho, H.-K. (2011). Plato's *Theaetetus* as an Indirect Dialogue: On the Difficulties in Writing Philosophy. *Soochow Journal of Philosophical Studies, 24*, 1-21. https://doi.org/10.29732/SJPS.201108.0001

Long, A. G. (2013). *Conversation and Self-Sufficiency in Plato.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199695355.001.0001

Loos, E., & Radicke, J. (2024). Using ChatGPT-3 as a Writing Tool: An Educational Assistant or a Moral Hazard? Current ChatGPT-3 Media Representations Compared to Plato's Critical Stance on Writing in *Phaedrus*. *AI and Ethics*. https://doi.org/10.1007/s43681-024-00470-1

Merriam-Webster. (2024). Generative AI. In *Merriam-Webster.com*. Retrieved 10 August 2024, from https://www.merriam-webster.com/dictionary/generative%20AI

Nvidia. (2024). Generative AI. In *Nvidia Glossary*. Retrieved 23 August 2024, from https://www.nvidia.com/en-us/glossary/generative-ai/

Plato. (1901). *Platonis Opera II: Parmenides, Philebus, Symposium, Phaedrus, Alcibiades I, II, Hipparchus, Amatores* (J. Burnet Ed.). Oxford University Press. https://doi.org/10.1093/actrade/9780198145417.book.1

———. (1902). *Platonis Opera IV: Clitopho, Respublica, Timaeus, Critias* (J. Burnet Ed.). Oxford University Press. https://doi.org/10.1093/actrade/9780198145448.book.1

———. (1995). *Platonis Opera I: Euthyphro, Apologia, Crito, Phaedo, Cratylus, Theaetetus, Sophista, Politicus* (E. A. Duke, W. F. Hicken, W. S. M. Nicoll, D. B. Robinson, and J. C. G. Stra-

chan Eds.). Oxford University Press. https://doi.org/10.1093/ac-trade/9780198145691.book.1

Roochnik, D. (2024). Socrates' Critique of Writing. *Society, 61*, 700-705. https://doi.org/10.1007/s12115-024-00968-8

Rouse, M. (2024). Generative AI. In *Techopedia.com*. Retrieved 24 August 2024, from https://www.techopedia.com/definition/34633/generative-ai

Rowe, C. (2007). *Plato and the Art of Philosophical Writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511482625

———. (2015). Methodologies for Reading Plato. In *Oxford Handbooks Online.* (www.oxfordhandbooks.com). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199935314.013.28

Trotz-Liboff, L. (2023). Writing as *Pharmakon* and the Limits of Law in Plato's *Statesman*, *Phaedrus*, and *Laws*. *Polis: The Journal for Ancient Greek and Roman Political Thought, 40* (3), 391-414. https://doi.org/10.1163/20512996-12340415