

We Asked ChatGPT About Art Theory. It Led Us Down a Rabbit Hole So Perplexing We Had to Ask Hal Foster for a Reality Check

A simple ChatGPT inquiry turned into a very strange ride.

Ben Davis, March 2, 2023



Image by DALL-E.

Another title for this article could have been “The Case Against ‘The Case Against Art History.’”

Recently, the news has been awash with stories about the very weird, surreal, sometimes sinister things that A.I. text generators are squirting out. Even so, Microsoft’s Bing search engine is incorporating the most famous of these, ChatGPT, thereby advancing at lightning speed to make “A.I. assistants” the standard for how people find knowledge on the internet.

Is this the future? I don’t know! But here’s a little art-history experiment gone wrong that illustrates some of the pitfalls.

Last week, my colleague Naomi Rea chatted me, asking if I had ever heard of an essay called “The Case Against Art History” by the theorist Hal Foster, from the Summer 1992 issue of *October*.

I looked it up and could find no such article.

It turned out that Naomi had been using ChatGPT exactly as Microsoft intends it to be used: as a research aid. In this case, she had been investigating whether there was any good writing on the concept of “category collapse” as it applied to contemporary art. The chatbot was extremely helpful. It quickly produced the following:

Yes, there are many essays and articles that explore the concept of category collapse in art history. Here are some examples:

1. *“Category Collapse and the Art Historical Canon” by Rebecca Zorach (Art Journal, Vol. 62, No. 4, 2003, pp. 76-87)*

This essay explores how category collapse can lead to the exclusion of certain artists and artworks from the art historical canon.

2. *“Re-Categorizing Art History” by Trevor Stark (Journal of Art Historiography, No. 19, 2018, pp. 1-15)*

This essay examines how the practice of categorizing art history can lead to oversimplification and argues for a more flexible approach to categorization.

3. *“The Case Against Art History” by Hal Foster (October, Vol. 61, 1992, pp. 3-22)*

This essay critiques the discipline of art history for its reliance on categorization and argues for a more interdisciplinary approach to the study of art.

4. *“The Problem of Category in Art History” by Whitney Davis (Art Bulletin, Vol. 83, No. 2, 2001, pp. 297-318)*

This essay discusses the limitations of categories in art history and explores alternative methods of organizing and interpreting art historical material.

Naomi followed up by asking it to summarize the Foster, to see if it might be useful. Again, the robot quickly supplied the goods via a concise and clear summary:

In “The Case Against Art History,” Hal Foster argues that the traditional discipline of art history is in crisis and needs to be rethought. He argues that the discipline’s focus on visual analysis and categorization has limited its ability to engage with the social and political contexts in which artworks are produced and received.

Foster critiques the idea that art history is a neutral, objective discipline and argues that it is shaped by power relations and ideological assumptions. He contends that the discipline is complicit in maintaining the status quo by privileging certain artists and artworks over others and by perpetuating myths about the autonomy of art.

Foster calls for a more interdisciplinary approach to the study of art that draws on methodologies and theories from other fields such as sociology, anthropology, and cultural studies. He advocates for a focus on the social and political dimensions of art, including issues of race, class, and gender.

Overall, Foster’s essay argues that the discipline of art history needs to move beyond the traditional emphasis on visual analysis and categorization and engage with the complex social and political contexts in which artworks are produced and received.

Pretty convincing! If only any of it was real. It’s all made up.

The title “The Case Against Art History” *sounds* like something I might have read in *October*. Hal Foster is real, as are Rebecca Zorach, Trevor Stark, and Whitney Davis. But the titles and articles are all made up—“hallucinations” of the system.

I have to admit, ChatGPT’s summary of “The Case Against Art History” sounds convincing on first scan. It feels like it summarizes something in the manner of a typical academic abstract. When I reread it, however, I realize that what it describes is not very specific, just something very general, like a baseline understanding of a social history of art.

But then, the citation was supposed to be from 1992. So a first take could easily be, “maybe that was a new idea back then...”

It’s as if the A.I. reduced the name “Hal Foster” to the lowest common denominator level of general knowledge that the internet might know about him—“often cited critical art theorist”—then speculated about what someone who fit that description might have contributed to a discussion of “category collapse” in art, the subject of Naomi’s query.

The other interesting thing about this episode is that, after we realized what had happened, Naomi looped back to ChatGPT to give it another chance. The chatbot was reasonably contrite:

I apologize for the confusion. It appears that “The Case Against Art History” is actually a chapter in Hal Foster’s book “Recodings: Art, Spectacle, Cultural Politics” (The New Press, 1985). I apologize for my error earlier.

Also not true.

Naomi followed up again:

I apologize for the confusion. You are correct that “The Case Against Art History” is not a chapter in Hal Foster’s book “Recodings: Art, Spectacle, Cultural Politics”. I apologize for the misinformation.

*I have double-checked and it appears that “The Case Against Art History” was originally published in the journal *October* in 1992. You should be able to find the article online through academic databases such as *Jstor*.*

Talk about a “category collapse”!

It feels like a parody of that scene from *2001: A Space Odyssey* where HAL 9000 refuses to open the pod bay doors: *I'm sorry, Naomi, but I'm going to have to insist that "The Case Against Art History" is a real citation... This mission is too important for me to allow you to jeopardize it...*

("Hal the computer's been cranking his stuff out since 1968," Foster joked when I wrote to confirm that "The Case Against Art History" was, indeed, fake. "He got rid of the meat Hal a long time ago.")

I played around with ChatGPT a bit further to get a sense of what it can tell me about art history. Sometimes—if, for instance, I ask it to "give me a list of citations about the influence of Artificial Intelligence on European Medieval Art"—it accurately tells me that this query makes no sense—but then provides a list of made-up references anyway:

I'm sorry, but there is no evidence to suggest that Artificial Intelligence (AI) has had any influence on European Medieval Art. AI technology did not exist during the medieval period, and there is no indication that it has had any retrospective impact on the art of that time.

However, if you are interested in exploring how AI is being used to study and analyze Medieval Art, I can suggest some references that discuss this topic:

- 1. "Medieval Art and Artificial Intelligence: A Review" by Heather E. Grossman, published in the journal "Digital Medievalist" (2019).*
- 2. "Computational Analysis of Medieval Art: Artificial Intelligence and Iconography" by Emilie E.S. Gordonker and Daan van Espen, published in the journal "The Art Bulletin" (2021).*
- 3. "AI and Art History: A Review of Current Research" by Carolyn Yerkes, published in the journal "Visual Resources" (2020).*
- 4. "Artificial Intelligence and the Study of Art: A Critical Review" by Hannah Turner and Florian Wiencek, published in the journal "Journal of Cultural Analytics" (2020).*

These articles explore the ways in which AI technology is being used to analyze and understand medieval art, including aspects such as

iconography, style, and historical context. However, it should be noted that AI is not influencing the creation of new works of medieval art. (I love the deadpan comedy of that final clarification—though arguably it is wrong given that DALL-E and its ilk can spit out infinite amounts of stuff in the style of “medieval art” now.)



“Illustration from a Medieval Book of Hours advertising fast food” by DALL-E.

When I ask a follow-up, specifying that the references now be actually “real,” my chatbot helper is again very helpful, but again just makes stuff up:

Certainly! Here are some real references that discuss the influence of Artificial Intelligence on Art:

1. *“The Art of Artificial Intelligence: GANs and the Next Step in Human-AI Collaboration”* by Ahmed Elgammal, published in the journal *“Leonardo”* (2018).
2. *“Art and Artificial Intelligence”* edited by Aude Oliva and Ahmed Elgammal, published by MIT Press (2020).
3. *“AI Art at Christie’s Sells for \$432,500”* by Kate Brown, published in *The Art Newspaper* (2018).
4. *“Can Computers Create Art?”* by Claudia M. Müller and Michael Leyton, published in the journal *“Leonardo”* (2019).

These references discuss the use of Artificial Intelligence in the creation and analysis of art, including works of art that have been generated entirely by AI algorithms. These articles are not specifically about European Medieval Art, but they do explore the intersection of AI and art more broadly.

Hilariously, one of these (the third) invokes a real colleague of mine, Kate Brown, who wrote no such thing.

I asked Naomi for her takeaways about the tool as a research aid. “ChatGPT is a useful tool but anyone using it should understand it is also unreliable,” she writes. “If you take what it says at face value, it could be a dumb and dangerous mistake.”

To be clear, sometimes it *does* generate real results. The technology is getting better by leaps and bounds, and this seems an obvious flaw to fix. But in the meantime, it’s bound to make everyone have to actually scrutinize any list of references with extra-paranoid suspicion. And you’re bound to have a lot more moments like the one I had with Naomi—someone asking you why they can’t find something that doesn’t exist.

The glitch seems to be a linear consequence of the fact that so-called Large-Language Models are about predicting what *sounds right*, based on its huge data sets. As a commenter put it in an already-months-old [post](#) about the fake citations problem: “It’s a language model, and not a knowledge model.”

In other words, this is an application for *sounding like an expert*, not for *being an expert*—which is just so, so emblematic of our whole moment, right? Instead of an engine of reliable knowledge, Silicon Valley has unleashed something that gives everyone the power to [fake it like Elizabeth Holmes](#).

