The Dance of the Doppelgängers: AI and the cultural heritage community

Susan Hazan
Digital Heritage, Israel
Maagal Beit Hamidrash 22, Jerusalem, Israel
Susan.hazan@gmail.com

There are severe issues that bring the impact of AI enabled systems into sharp focus for the cultural heritage sector, especially when these platforms are open, extremely tempting, and now accessible for everyone who may wish to create art or write a novel at the keyboard. Creating the verbal cue and waiting for the sublimity of every desire is mesmerizing, enthralling, even addicting but the impact on human creativity somehow feels like an accident waiting to happen. This profuse human/machine creativity resulting in AI artifacts sets up enormous challenges for all players in the creative industries, from the struggling artist or designer; to the public-funded institutions such as museums who will be wandering the labyrinth of copyright policies, trying to inscribe the metadata of the provenance of AI production. The paper will discuss Chat GPT and present AI imagery as doppelgängers that seemingly emerge out of an uncanny valley. The process is almost immediate with DALL-E2, Midjourney, and Stable Diffusion all providing instant results at the drop of a prompt. With a series of case studies, we will discuss the challenging copyright and provenance issues that are taking place at a dizzying pace as a sector that is striving to come to terms with the highly problematic issues of authenticity and ownership.


Note the level columns on last page.

1. INTRODUCTION

AI is currently attracting a lot of interest across all aspects of popular culture, in novels, the media and social networks but the field is relatively new, and it is interesting at this point to consider the parameters of AI’s capabilities, and the robustness of its competencies. There seems to be a severely hyped promotion of it’s genuine potential, as if AI had hired an amazing PR manager and over-promised what these systems can actually deliver. This paper will showcase the impact of AI on the cultural heritage community and the creative industries; two sectors who value human creativity and prioritise the unique and the authentic. AI artifacts, on the other hand authored by DALL-E2, Midjourney, and Stable Diffusion all provide instant results at the drop of a prompt while challenging notions of authenticity, trust, and ownership. The fantastic imagery that emerges from these machine/human interactions has vast potential for artists, photographers, and graphic designers as they create their doppelgängers that seemingly emerge out of an uncanny valley. However, embedded in all this magic are serious issues that reveal the challenges for these sectors, especially when the public begins to explore open access platforms for themselves, when everyone can be sublimely creative, even without any kind of previous training. Creating the verbal cue and waiting for the outcome of every desire is mesmerizing, enthralling, even addicting but the impact on human creativity somehow feels like an accident waiting to happen. The profound human/machine creativity that spawns AI artifacts sets up enormous challenges for all players in the cultural heritage and creative industries, from the struggling artist or designer; to the public-funded institutions such as museums who will be wandering the labyrinth of copyright policies and trying to inscribe the metadata of the provenance of AI production. This paper will discuss the challenging copyright and provenance issues that are emerging at a dizzying pace in the sectors that are striving to come to terms with highly problematic authenticity and ownership issues. AI artifacts are spawning at an alarming rate in the public realm and are besieging the cultural heritage community with all kinds of visual and textural abominations.
We will start with ChatGPT which stands for generative pretrained transformer. Developed by OpenAI, and released for testing to the public November 30, 2022, ChatGPT is (currently a free) chatbot that can generate an answer to almost any question requested. But who am I to define the bot – why don’t we let it describe itself. I originally asked it to do exactly that and, within seconds I received a pleasant response. Perhaps trickier was to challenge the system to find out how much it is aware of what it can't do, for example, or if it can sense the limits of its own scope. Ironically it answered in the third person.

Describe what ChatGPT cannot do?
As an AI language model, there are several things that ChatGPT may not be able to do or limitations that it may have. Some of them are:

1. It cannot fully understand human emotions and feelings: ChatGPT is designed to analyze and process language, but it does not have emotions or feelings. It can simulate a conversation and provide responses based on the input it receives, but it cannot fully empathize or understand the emotional state of the user.
2. It cannot perform physical tasks: ChatGPT is a software-based AI language model that exists only in the digital world. It cannot perform any physical tasks or interact with the physical world directly.
3. It cannot replace human interaction completely: While ChatGPT can simulate a conversation and provide answers to user queries, it cannot replace human interaction completely. There are certain situations where human interaction is required, especially when it comes to complex emotional or social interactions.
4. It cannot generate completely original content: ChatGPT is designed to generate responses based on the input it receives. While it can generate new text based on its training data, it cannot generate completely original content that is not based on any pre-existing data.
5. It cannot make moral or ethical judgments: ChatGPT is designed to provide responses based on the input it receives, but it cannot make moral or ethical judgments. It can provide information and answers based on its training data, but it cannot determine what is right or wrong.

I found that a candid response and I continued to explore its capacity to deliver the goods. Each time I was surprised at its inventive responses but also deeply troubled by some of its somewhat dubious results. I challenged its capabilities in as many ways I could think of - it wrote me a poem, coded some HTML for me – and yes it worked; it wrote the music score for guitar for our national anthem, and it did some modest research for me. It supplied the goods in much the same way we would expect a Google search to accomplish – only presented more intuitively. I played around with it with great enthusiasm and even satisfaction until – it smacked me right in the face. I asked it to write an essay on AI in contemporary culture in the style of Dr Susan Hazan. The results were shocking. It blew my head off. I felt that something had got into my mind and was scraping my brain. This was my prompt...

Write an essay on AI in contemporary culture in the style of Dr Susan Hazan. This is what came out...

Artificial Intelligence (AI) has become a pervasive and defining aspect of contemporary culture. The rapid advancements in technology and its widespread adoption have made AI an integral part of our daily lives, impacting everything from the way we communicate to the way we do business. Dr Susan Hazan, a leading researcher in the field of AI and Culture, has noted that AI is increasingly being used to create cultural products and experiences, such as music, visual arts, and virtual reality. AI algorithms can generate new forms of expression that are changing our understanding of what it means to be creative. For example, AI-generated music is blurring the lines between the work of human musicians and machines, challenging traditional notions of authorship and creativity.

Moreover, AI is changing the way we consume culture. Hazan points out that AI-powered recommendation systems are shaping our taste and influencing what we watch, listen to, and read. These algorithms, based on our individual preferences and consumption habits, create personalized cultural experiences that are rapidly becoming the norm. The result is a fragmentation of cultural taste, where people are less exposed to a shared cultural repertoire and more likely to consume content that aligns with their existing preferences.

Another area in which AI is having a significant impact on contemporary culture is in the realm of social media. Hazan argues that AI algorithms are shaping our online interactions, influencing what we see, who we connect with, and how we express ourselves. These algorithms are designed to maximize user engagement and generate profits for the companies that own them, which has led to a culture of outrage, clickbait, and misinformation. Hazan suggests that AI has the potential to both liberate and constrain human expression and calls for greater transparency and accountability in the design and use of AI systems.
In conclusion, AI is a transformative force in contemporary culture, changing the way we create, consume, and interact with cultural products and experiences. While AI has the potential to enhance our lives and expand our cultural horizons, it is important to understand its limitations and to be aware of the ways in which it is shaping our culture and our society. As Hazan argues, the impact of AI on culture and society is complex, and it is essential to engage in critical and informed discourse about its use and implications.

Most of the Hazan ‘points out’, ‘argues’ and ‘suggests’ were surreptitious ways of getting round the direct quote – of course I had never ever said any of this – I have barely said these things out aloud, let alone in public. But here this ‘thing’ was crawling around my published chapters and papers; perhaps my Twitter feed, and other corners of the Internet and presuming to come out with the conclusions I would come to myself based on what it had learned about me. My ideas had become a hotchpotch of presumptions that, when stringed together spewed out a convincing doppelgänger of me. Hopefully, sometime in the future, once I present my ideas at enough conferences and publish enough chapters and articles like this one, I might claim the title ‘leading researcher in the field of AI and Culture’. Did it slip in that compliment to subdue me and lull me into a sense of entitlement? Sadly, as Janelle Shane, optics research scientist and artificial intelligence researcher points out ‘today’s AI is much closer in brainpower to an earthworm than to a human. It can pattern-match but doesn’t understand what it’s doing (Shane 2019).’ Additionally, without the context it wouldn’t even know what a compliment was—even if it tried.

In a later search I asked ChatGPT who is Dr. Susan Hazan’ and the results were truly uncanny and included ....

Throughout her career, Dr. Hazan has been recognized for her contributions to the field of digital cultural heritage and museums. In 2017, she was awarded the Chevalier des Arts et des Lettres by the French Ministry of Culture and Communication, and in 2021 she was named one of the Top 20 Most Influential Women in the Arts and Culture Sector by Artlyst.

I was so troubled by this I actually went to the Artlyst site to see if perhaps I had received this award, and nobody had told me – sadly not. So how did this earth worm get to presume me? Trained on unimaginable amounts of text, ChatGPT strings together words to produce reasonably coherent sentences based on statistical probability of which words follow other words. Since its training was completed in 2021, it has limited knowledge of the world and events after 2021 which is a sure way to create false data or even biased content. According to a recent message from Open AI, the initial training period data submitted through the API is no longer used for model training or other service improvements, unless you explicitly opt in. At the same time, my secret hope is that ChatGPT is reading cryptic patterns on the Internet and its presumptions about my receiving the Chevalier des Arts et des Lettres by the French Ministry of Culture and Communication might come true in the future. ChatGTP no longer asks me which Susan Hazan I am referring to, as it did on my first query, it now knows! I suspect that my own interaction with ChatGPT is further burnishing my fame as AI expert as reported by my recent prompts and possibly spewing these new ‘facts’ as truth back into the system for posterity. For the record, until now I have not yet published anything specifically on AI and digital culture.

Stable Diffusion is an artificial intelligence software product, released in August 2022 by a company called Stability AI. Stable Diffusion draws on unauthorized copies of millions—and possibly billions—of copyrighted images, all harvested without the knowledge or consent of the artists. According to Matthew Butterick, after scraping the billions of images—“without the consent of the original artists—Stable Diffusion relies on a mathematical process called diffusion to store compressed copies of these training images, which in turn are recombin ed to derive other images. It is, in short, a 21st-century collage tool”. Both ChatGPT and Stable Diffusion create new content, including audio, code, images, text, simulations, and soon to come also videos (ChatGPT4) based on previous content to produce unique versions of a mashup of prior art. Both ChatGPT and DALL·E received billions in funding but while ChatGPT was launched as a free tool it’s now getting monetized via a premium subscription model sometimes known as a freemium model. Since the natural language processing model is still in its research and "learning" preview phase, people can use it for free; all you need is to register for a free Open AI account, though there is an option to upgrade to a paid membership.

2. WHERE IS IT TAKING US: BIAS, COPYRIGHT AND AUTHENTICITY

2.1 Bias

In what is referred to as the Fourth Industrial Revolution (or Industry 4.0) we carry on with our lives often oblivious to the systems that make decisions for us, or about us through machine-readable solutions which bypass human
intervention. However, rather than being innocent the omnipresent machine-to-machine automated processes may be infested with their own predispositions; those coded by humans who, perhaps transpose their own biases or misconceptions through pre-trained data with a penchant for re-enacting privileged metadata solutions already in place (Hazan, 2020). This becomes magnified to even a greater magnitude in big data when massive datasets are automatically compiled; pre-polluted.

Big data magnifies these kinds of biases exponentially because, at the end of the day, it is always human agency that has been doing the irreversible hard coding. Not all biases are malicious of course, they can be often unconscious or unintended but still very much present in the data. The realization that a system or big data set might not be objective must always be taken into consideration and large automated systems should be approached with caution. In the city of Rotterdam in 2021, in an effort to rank people based on their risk for fraud a system used by the Dutch city of Rotterdam, thousands of people across the city were investigated by welfare fraud officers who sought out individuals suspected of abusing the system. The machine learning algorithm generated a risk score for each of Rotterdam’s roughly 30,000 welfare recipients, based on specific parameters—being a parent, a woman, young, not fluent in Dutch, or struggling to find work—all increasing someone’s risk score. Acting on these suspicions driven by the algorithm the city essentially discriminated against individuals through no fault of their own. Around the same time the Netherlands was struggling with a similar nationwide machine learning scandal when more than 20,000 families were wrongly accused of childcare benefit fraud after a machine learning system produce false positives. This resulted in forced evictions and decimated entire families. The entire Dutch government resigned in response in January 2021. The city’s decision to pause its use of the welfare algorithm in 2021 came after an investigation by the Rotterdam Court of Audit on the development and use of algorithms in the city.

So, can we trust AI? As Janelle Shane reminds us, treating a decision as impartial just because it came from an AI is known sometimes as mathwashing or bias laundering. Shane knows a thing or two about training and has spent many hours training AI systems which she recounts in her hilarious book where according to her experience, data is at times, unruly, surprising, and often unpredictable. According to Michael Schmidt, Chief Technology Officer, DataRobot, the largest source of bias in an AI system is the data it was trained on. That data might have historical patterns of bias encoded in its outcomes. Ultimately, machine learning gains knowledge from data, but that data comes from us—our decisions and systems (Schmidt 2023). How does this impact our results when we dabble with an image generating system, DALL-E for example. One red team member told WIRED that eight out of eight attempts to generate images with words like “a man sitting in a prison cell” or “a photo of an angry man” returned images of men of colour. “There were a lot of non-white people whenever there was a negative adjective associated with the person,” says Maarten Sap, an external red team member who researches stereotypes and reasoning in AI models. “Enough risks were found that maybe it shouldn't generate people or anything photorealistic.” This of course has since been rescinded with AI models creating impressively photorealistic images of all kinds of people. Sadly, due to an inbuilt latency problem, sometimes people appear with an extra finger or a weird limb sticking out in a strange part of the body; all developments that are atomically ludicrous and sometime quite monstrous.

In my own experience with DALL-E and Midjourney the systems warned me several times for using terms that were deemed inappropriate by their filters. In a search for background material for Sir Hans Sloane and the slave trade for example, my request was strangely banned when I prompted an image using the term ‘slave trade’.

---

**Figure 1**: Response to slave prompt from MidJourney

**Figure 2**: Response to slave prompt in DALL-E2

I discovered a similar ban when prompting the systems to create images of the Holocaust for my academic research which truly startled me. I felt that I had broken some sort of unspoken rule. Evidently,
I had. According to the document authored by Open AI ethics and policy researchers DALL·E 2 was trained using a combination of photos scraped from the internet and acquired from licensed sources, Open AI was making efforts to mitigate toxicity or the spread of disinformation, applying text filters to the image generator and removing some images that were sexually explicit or gory. Evidently the terms ‘slave’ and ‘Holocaust’ are protected terms and we need to tread gently when invoking these kinds of images. There is a comprehensive list of banned prompts online for the curious.

2.2 Copyright

Perhaps the toughest challenge for a museum is the question of copyright. There are two main issues around copyright and AI image generators. The first is whether the images that were used to train the software have been licensed and whether permission was granted. The second is the issue of who owns the copyright of the generated image. The various image generators trained their models on millions of pictures scraped off the web but more often without telling their creators or seeking their consent. This leaves the field wide open for rampant exploitation and in my opinion is basically an accident waiting to happen. Legal proceedings are already underway. In the United States, a class action was launched by Mathew Butterick who is representing people from all over the world—specifically writers, artists, programmers, and other creators—who are concerned about AI systems being trained on vast amounts of copyrighted work with no consent, no credit, and no compensation. In the United Kingdom, Getty Images, a supplier of commercial photographs, is suing the company behind Stable Diffusion. Getty Images claims Stability AI ‘unlawfully’ scraped millions of images from its site, which in their opinion is a significant escalation in the developing legal battles between generative AI firms and content creators. Shutterstock is now removing AI-generated Images. The flood gates are already open, and AI generated images are besieging the cultural heritage sector. The battles have just begun. The move towards litigation is gathering momentum. Midjourney’s terms of service now includes a DMCA takedown policy, allowing artists to request their work to be removed from the set if they believe copyright infringement to be evident. GitHub offers a Guide to the Digital Millennium Copyright Act, commonly known as the “DMCA as a pathway to those bewildered.

To explore the potential of AI text to image generators I created an exhibition using the prompt ‘oil paintings, based on the style of Rembrandt’. My goal was to test the potential of the interface and I was astounded when the magical creatures started appearing on my screen. The Uncanny Animal Musicians – an AI Exhibition soon took shape, populated with fantastical animal musicians. My family loved them, and they are now hung next to the piano for inspiration and have their very own website.

Figure 3: Screenshot from the online exhibition – The Uncanny Animal Musicians

My doubts about my own integrity first seeped in when I went to the shop when I printed the digital creations on canvas. The person who handed me my newly minted creations asked me who had painted them and, without blinking an eyelid, I answered ‘Rembrandt’. Arguably the main problem with AI image generators concerns the general lack of regulation around the technology but innovative solutions are already in the making. According to Joseph Foley, a system called Glaze, ‘hides’ the artistic style by applying ‘style cloaks’ to their work before they upload them to the web which misleads the generative AI model when it tries to mimic an artist's style, causing it to associate the artwork with something else. With the demand that has become so craven as the days go by, more and more AI blocking and cloaking systems will be hitting the market.

2.3 Authenticity

In addition to the challenges posed by the copyright issues we also have troubling issues with notions of authenticity. Early 2023, ChatGPT sat for, and passed all three parts of the U.S. Medical Licensing Examination, although just barely, as part of a research experiment. Instead of the usual 2 years of preparation medical students usually undertake, ChatGPT performed well enough to pass the exam despite never having been trained on a medical dataset. Although this was an experiment carried out in controlled conditions, the potential for fraud is disconcerting. Where human creativity can be monetised, AI may present many similar challenges. In early 2023, a major sci-fi magazine halted submissions due to flood of stories written by AI.
chatbots. Clarkesworld pays for submissions that are accepted where the fee is based on 12¢ per word, meaning that longer submissions, if accepted, could pay out a reasonable cash sum. Although the publisher had been dealing with plagiarism for many years, with copy lifted of the Internet and rehashed into a new sci-fi story, according to Neil Clarke, editor-in-chief of Clarkesworld "towards the end of 2022, there was major spike in plagiarism as 'AI' chatbots started gaining some attention, putting a new tool in their arsenal, and encouraging more to give this 'side hustle' a try. It quickly got out of hand." Clarke posted a graph from the year showing how the number of articles spiked significantly. This caused Clarke to close submissions temporarily and Clarkesworld announced that is "not considering stories written, co-written, or assisted by AI at this time."

![Figure 4: Clarkesworld's AI submissions](image)

Arguably the main problem with AI image generators concerns the general lack of regulation around the technology but things are moving fast. In the case of OpenAI’s ChatGPT, the company has tried to mitigate against such developments with a content moderation tool, one that's designed to flag and block unsafe and illegal information as mentioned above. Popular programs like DALL-E, Midjourney, and Stable Diffusion already have policies in place that prevent consumers from using their products in certain ways, like appropriating images of celebrities and politicians. But there is still however, a vast grey area that needs plugging. In October 2022, a popular photographer decided to be more honest about his creativity. Jos Avery of Averyseasonart was amazed at the popularity of his photographs which had, according to him - wildly exceeded his expectations. This uneasiness led him to contact Ars Technica, a technology publication with the statement “Probably 95 percent-plus of the followers don’t realize [the images are A.I.-generated],” he told Ars Technica, “I'd like to come clean.” How many more doppelgängers are lurking out there waiting for their embarrassing disclosure?

3. **NEWLY DISCOVERED ARTIFACTS**

FOR IMMEDIATE RELEASE
Recently Discovered Rare Fashion Items in Movie Star’s Private Collection

Los Angeles, California - In a stunning discovery, a collection of rare fashion items has been found in the private collection of a famous movie star. The collection includes an array of glass art deco handbags that are highly sought after by collectors around the world.

The collection was uncovered by a team of art and fashion experts who were given access to the movie star’s private collection. They were amazed to find such a vast collection of fashion items that had never been seen before.

The glass art deco handbags in the collection are some of the rarest and most sought-after fashion items in the world. They were created by some of the most well-known designers of the time and were highly coveted by Hollywood stars in the 1920s and 1930s.

In addition to the handbags, the collection also includes a number of stunning art deco jewellery pieces and vintage clothing items. The items in the collection have been carefully preserved and are in excellent condition, making them even more valuable to collectors.

The discovery of this collection is a major event in the fashion world and is sure to generate a great deal of excitement among collectors and fashion enthusiasts. The collection will be showcased in an upcoming exhibition, where members of the public will have the opportunity to view and appreciate these stunning pieces of fashion history.

Some of the items in the collection have already been auctioned off, fetching record prices at auction houses. A glass art deco handbag designed by Rene Lalique, for example, sold for $75,000, while a vintage evening gown worn by the movie star at a Hollywood gala fetched over $100,000.

The identity of the movie star has not been revealed, adding to the intrigue and excitement around the collection. It is believed that the star was a prominent figure in Hollywood during the golden age of cinema and had a keen eye for fashion and design.

The exhibition showcasing the collection is set to open at a prestigious museum in Los Angeles next month and is expected to draw a large
crowd of fashion enthusiasts, collectors, and movie fans from around the world.

For more information about the collection and the upcoming exhibition, please contact [contact person] at [contact email/phone].
Contact: [contact person] Email: [contact email] Phone: [contact phone]

And just in case you had any doubt what these art deco glass handbags looks like, you can see the ‘evidence’ with your own eyes.

Figure 5: Art Deco glass bags by Midjourney

Of course, none of this is real! All pure fabrication. I had requested ChatGTP to write me a scoop – a press release – with recently discovered rare fashion items in a movie stars private collection of glass, art deco handbags, and then asked Midjourney to provide the visual proof. I just had to fill in the contact details. Clearly art fraud and similar crimes are not new, but what is new is the ease of production now sitting at our fingertips; throw in a little creativity, and a little audacity and voila a new reality is born. Of course, this has always been the way with artists. Creating new realities with the paint brush, camera, or chisel, and now with digital platforms.

Jack Wylder has embraced image generators with enthusiasm, even publishing an entire book on the perfect prompt. According to Wylder …

I’ve heard many artists lamenting that this is the beginning of the end for art done by humans, but I disagree – it’s the start of a whole new way of creating art, no different that when cameras were first invented, or Photoshop first became available. The Programs can make beautiful art, but it is the spark of human creativity that organizes and makes sense of output; that is the irreplaceable creativity we bring to the partnership and why our positions as artists aren’t going anywhere. This new technology isn’t getting rid of artist – it’s creating thousands of new ones! I am excited to see what they do with it (Wylder 2023 p. 262).

3. CONCLUSION

I would like to temper my angst I have shared here with you with a few words of optimism. The image generators do produce fascinating photorealistic images, like Jos Avery’s pseudo works that are still remarkable even after he came clean and reminded us that these ‘people’ had never breathed a drop of oxygen in their lives. With life photography, however, there is little to be concerned about neither with genres such as journalism nor documentary. Photography that reflects our daily lives and records those special occasions that needed to be remembered is irreplicable. Neither Midjourney nor DALL-E can prompt this kind of photography with an AI generated image. Nor can ChatGTP inject the depth of human empathy and the intuitive contextualisation from a curt text prompt beyond the compassion of an earth worm or the capacity of a 9-year-old child. At the same time, and as Wylder suggest generative AI opens new skillsets and potentially new niche professions as all artistic production, is built upon layers and layers of prior creativity and culture – and as they say – we are all standing on the shoulders of giants.

Still, it is worth keeping our eyes and ears open for these sneaky doppelgängers when we confront new forms of creativity. You can never know what those algorithms have been doing behind out backs!

4. BIBLIOGRAPHY


Shane, J. (2019) You Look Like a Thing and I Love You: How Artificial Intelligence Works and Why It’s Making the World a Weirder Place, Headline Publishing Group, UK

