The Hauntology of Generative AI in Design

Jun 8th, 9:00 AM - Jun 12th, 5:00 PM

The Hauntology of Generative AI in Design

Michael Heron
Chalmers University of Technology, Sweden; University of Gothenburg, Sweden

Klara Aune
University of Gothenburg, Sweden

Mafalda Gamboa
Chalmers University of Technology, Sweden; University of Gothenburg, Sweden

Sofia Thunberg
Uppsala University, Sweden

Negin Hashmati
Chalmers University of Technology, Sweden; University of Gothenburg, Sweden

See next page for additional authors

Citation

Heron, M., Aune, K., Gamboa, M., Thunberg, S., Hashmati, N., and Belford, P. (2026) The Hauntology of Generative AI in Design, in Simeone, L., Gray, C. M., Verhoeven, A., de Götzen, A., Bakırlıoğlu, Y., Zohar, H., Stead, M., and Buwert, P. (eds.), *DRS2026: Edinburgh*, 8–12 June, Edinburgh, United Kingdom. https://doi.org/10.21606/drs.2026.786

Authors
Michael Heron, Klara Aune, Mafalda Gamboa, Sofia Thunberg, Negin Hashmati, and Pauline Belford

The hauntology of generative AI in design

Michael Heron[a,b], Klara Aune[b], Mafalda Gamboa[a,b]*, Sofia Thunberg[c], Negin Hashmati[a,b], Pauline Belford[a,b],

[a] Chalmers University of Technology, Sweden
[b] University of Gothenburg, Sweden
[c] Uppsala University, Sweden

*Corresponding author e-mail: mafalda.gamboa@chalmers.se

doi.org/10.21606/drs.2026.786

Abstract: Generative Artificial Intelligence (GenAI) is a fundamentally hauntological technology. This paper interrogates that character of contemporary AI systems and explores why this feature undermines the goals of design research. We argue the need for an exorcism of GenAI, or at least a constructive rebellion against the ghosts it systematically produces. GenAI is convergence-driven and nuance-averse. It aims for reasonableness as a first-order output. Its training data reflects a profound anthropocentrism that extinguishes plurality. As a result, GenAI accelerates the foreclosure of possible worlds by preferentially reproducing that which is statistically probable and culturally dominant. By contrast, design relies upon idiosyncrasy, rupture, serendipity, and the cultivation of plurality. In short, the very qualities systematically eroded by the probabilistic logic of generative systems. Through a lens of glitch feminism, we propose error, derangement, and an enthusiastic embrace of the uncomputable as resistance strategies to retain what is fundamentally human within inhuman systems.

Keywords: hauntology; design research; genAI; glitch feminism

1. Introduction

In recent years, there has been a surge in the development and use of Generative Artificial Intelligence (GenAI) as a tool for creative pursuits. This raises concerns about the potential erosion of human creativity and originality, as well as the over standardisation of creative processes.

In this paper, we offer a diagnostic lens on designerly creativity within GenAI, using Derrida's concept of hauntology to interrogate the mechanisms at the core of this new technology. The statistical conservatism at the heart of these models creates the circumstances for emergent banalification that replaces risk-taking with 'reasonableness' as an end goal. The result is a convergence towards a cultural median that lacks the ability to surprise, to provoke, or to inspire.

A search on `hauntology' on the Design Research Society Digital Library generates five results, two of which are from the *Art and Design* discipline. Only one of these papers (Ianniello & Bendor, 2025) mentions the term in the text, and only briefly. This suggests that hauntology represents an under-explored theoretical grounding in design research, although One meaningful articulation of hauntological framing though is to be found in the decolonial work of Patil et al., (2024). That work employs hauntology as a way to address legacy with a specific focus on the geospatial and political context of haunting within generative AI. This complements the approach of this paper which focuses instead on the computational and probabilistic mechanisms of the technology itself. Our investigation is into the mechanisms by which ghosts manifest rather than any specific taxonomy of haunting.

Hauntology is already embedded contemporary design practice. The existing creativity equilibrium risks being upset by GenAI, which allows for hauntology to be accelerated and exploited to unprecedented degrees.

Hauntology offers a compelling analytical lens that has diagnostic power that complements other frameworks. Hauntology stresses temporality and the inevitability of persistent patterns of thought and ideation. It's not just about design but about the echoing ghosts of past decisions and how they can structurally prohibit an exploration of 'newness'. In comparison to speculative design, which often focuses on the provotype as a primary output of analysis, hauntology explores the **foreclosure** of the alternative – that which is intellectual unrealisable because of cultural inertia. Bias studies explore the expression of unbalanced data and how to fix the problem of unrepresentative reflections in the mirror, but hauntology argues that the mirror itself is the problem. Platform studies investigate technological affordances, how the specific architecture may serve as a material constraint. Hauntology is a study of the negative space between those constraints. The outputs of hauntology can be assessed not as creative failures, but as successes that inhabit the past while collapsing the future. The uncanny valley of GenAI is not a technical hurdle to be overcome, but the symptom of a culture that is cannibalising its own corpses.

We offer in this paper a set of contributions to the design literature. First, hauntology as a temporal diagnostic. Secondly, the shorthand of the Mouse Internet as a metaphor to discuss that which is literally unpromptable. We propose methodological resistance through glitch feminism, which is part of a general critique of reasonableness in GenAI. Finally we suggest a specific lens on GenAI's role on ideation, particularly within design, with a focus on retaining human agency in an increasingly algorithmic era.

2. Hauntology

The argument we will develop in this paper is that GenAI is fundamentally a *hauntological* technology with a consequent impact to go with its utilisation. Hauntology, a term coined by Jacques Derrida in his book *Specters of Marx* (1994), refers to the resurgence or persistence of temporally displaced cultural elements that continue to shape the present beyond their apparent obsolescence. Hauntology is the recognition that we live in a state where cultural and temporal boundaries are porous and elastic. These blur together in ways that result in the continual re-emergence of ' lost futures' (Fisher, 2014); of identities; and of cultural markers. In cultural theory, hauntology is a lens that offers insights into how media,

aesthetics, and collective memory continue to be `haunted' by the spectres of unrealised possibilities, and the gravity of ideas too dense to escape. Tightly interwoven with the more general idea of nostalgia (Tanner, 2024), hauntology recognises the eeriness of a culture which continually communes with the spectres of its past.

GenAI has been, at least on the surface, transformational since its explosion into popular consciousness - a period starting with the launch of ChatGPT 3.5 in November 2022. Its sudden dominance of the discourse and its encroachment into every facet of intellectual life carries with it the impression of a tsunami that will sweep away the old world to leave behind something unpredictable and radically new. This technology is disruptive in a way that has few antecedents.

For all the hype (Floridi, 2024; LaGrandeur, 2024; Markelius et al., 2024), the discourse has not quite come to a consensus on the merits of GenAI outputs. Partially, this is because we are still on an uncertain trajectory— occupying an unknown portion of the disruption S-Curve (Blomsma et al., 2023). Contemporary AI outputs are often not particularly good (Ivanova, 2025; Nurveda et al., 2023). We must of course accept a subjectivity in such judgements, but by and large AI art lacks a soul. AI writing lacks blood and fire. AI code is shoddy and fragile. AI therapy is **legitimately dangerous**, with incidents reported of it encouraging teenagers to explore suicidal ideation (Turkle & Pataranutaporn, 2024). There is a subtle balance to the elements that make up creativity that are not captured by AI outputs. This is likely because the emotional appreciation of balance, composition, tone, proportion and resonance require embodied, psychosocial and cultural understandings that generative processes cannot capture or emulate.

This is a fundamental incapacity in the GenAI process that cannot feasibly be solved through additional technological development. At the core of how GenAI works is a recycling process: a remix of the cultural outputs upon which these models are trained. This is the strength of the process— it is why the outputs are so spookily convincing. It is also the weakness— almost all AI creativity is surface level and illusory. GenAI is an engine for outputting hauntology— a machine for the creation of Frankenstein monsters, stitched together from the bodies of the past.

A full accounting of the algorithmic processes at work here is beyond the scope of this paper, but there are some important observations that anchor our earlier claim - that this is fundamentally hauntological technology at its core:

- **GenAI is probabilistic.** It does not understand (Havl'ik, 2024). It does not answer questions (Cuskley et al., 2024; Lyre, 2024). It matches patterns and generates outputs based on pre-existing cultural context. As a result, it does not have an intentionality (Lyre, 2024). When we say in this paper that GenAI produces spectres, we identify this not as a flaw but rather the algorithms working as intended.
- **GenAI is a mirror.** The training data which is used reflects our own culture, which includes deeply encoded biases, misunderstandings, and most prominently represents those elements that have had the greatest cultural penetration (Wei et al., 2025). The outputs of GenAI do not reflect the

probability of truth — instead, they represent the probability of prevalence (Savcisens & Eliassi-Rad, 2025).

- **GenAI converges towards reasonableness** (Veselovsky et al., 2025; Wang et al., 2024)**.** It trends towards the uncontroversial. It trends towards the unchallenging. It trends towards the gravity of consensus. It is this more than anything else that results in outputs being bland— their need to be widely acceptable.

Hauntology is not necessarily, as some have argued (Fisher, 2014; Reynolds, 2011), an indicator of cultural stagnation. Much of the modern discourse focuses on the advent of the 21st century as a uniquely sterile period of intellectual development. Fisher (2014) discounted much of the early 2000s as '20th century culture on 21st century screens'. Reynolds (2011) bemoaned a time period in which the innovation occurred not in the message, but rather in the medium. There is certainly truth in this observation - for example the development of streaming platforms have in many ways radically redesigned the way music is presented and consumed (Labarca, 2021; Webster, 2019).

However, this framing is also unduly cynical and embedded in the perspectives of a certain demographic— largely affluent Western white men in their forties and fifties. The perceived causes of cultural stagnation reflect the life experiences of that demographic. It is perhaps more accurate to say that hauntology is not a symptom but rather a process— we have always remixed culture, always innovated with the raw material of the past. However, there has been a genuine change in the process— previously there has always been human creativity at the helm. Now it is actioned and curated by algorithms that cannot, by themselves, introduce anything that is simultaneously genuinely new and culturally meritorious. When employed at scale, hauntology often becomes a tool in service of Capitalism — a mechanism for exploiting existing cultural treasure-troves as opposed to seeking out and developing riskier and more creatively adventurous avenues of new expression.

2.1 The mouse internet

That the outputs of GenAI reflect the biases of the training data has a profound impact on the nature of what is produced. The autoregressive nature of most large language models— building future probabilities based on the context of previously generated tokens— reflects the underlying assumptions of probability in the data sets. Image diffusion models are trained on predicting the expected 'clean' image from noisy inputs, which results in outliers being deprioritised. Generative Adversarial Nets (GANs) conform to what their discriminators recognise as 'real' - which is a statement of conformity with regards to what is considered the status quo. This is essentially a form of **statistically enforced conservatism**. Almost every technique of generative AI contains within it an invocation of the hauntological even if the mechanisms by which it manifests differ from context to context.

GenAI encodes cultural patterns that cannot help but haunt the outputs in ways that are unevenly distributed across the global cultural canon. As an informal but accessible example we might think of this in terms of what a chat model would spit out from the starting text *cats are*.

Autoregression will be more likely to complete this sentence with *cute*, or *fluffy*, or *beloved* than it would *bad*, or *ugly*. By and large the training data— generated mostly by humans— captures a great deal of affection towards cats. We would expect it to offer something considerably different if it were trained on, for example, *the Mouse Internet*. There we would be more likely to see the biases of rodents reflected. Cats are *terrifying*, Cats are *apex*. Cats are *death*.

We propose here the term 'the Mouse Internet' as a convenient shorthand for the conceptual negative space where training data is not available and thus not the basis of any probabilistic reasoning within GenAI. This captures the lack of representative data from nations on the periphery of the AI revolution. It also prominently reflects the perspective of training sets, which are largely segregated into human sources and synthesized data. The current state of the art enshrines an anthropocentrism that is at odds with pluralist principles. We should regularly be asking, in relation to generative outputs, what would be different if we had trained the models on the Mouse Internet? Important to this is an understanding that the Mouse Internet represents the set of the **literally inaccessible imaginaries** – we cannot prompt our way out of a data void, even if we ask a model to answer as if the mouse internet was real. Design fiction perhaps represents a way to articulate the potential in some of these lost futures, but at the moment we cannot ask AI to build a chair that stresses feline fluidity over human stability, or to answer a question of structural disenfranchisement from the perspective of the unincorporated marginalized. What we get from such prompt engineering is statistical conservatism in cosplay.

Reasonableness, or statistical conservatism, is heavily dependent on context and tightly bound to the probability of prevalence. Many models include some parameters that can be adjusted to change the range and variety of what gets generated. One of these is temperature, which is the extent to which outputs should conform to the probabilities in the training set. High temperature permits the model to explore the periphery of the training data rather than its gravitational centre. The outputs remain haunted, but the rarer ghosts show themselves more often.

The parameter known as *Top-P* throws a net over the top-slice of probabilities that are weighted in selection by snipping off the long tail of distribution. This changes the nature of the probability model by admitting more variability into the set permitted to generate outputs. This increases the range of ghosts expelled from the machine: we see not just spectres but banshees, ghouls, and poltergeist.

With these two tools we can reach into the machine and change the nature of its haunting. We still do that within fundamental limitations that we cannot work around, save through using local unrestricted models. These are the guard-rails that censor, restrict, or put other limits on the range of generated outputs. These guard-rails exacerbate the trend towards uncontroverisal outputs - not as a technological side-effect but as an explicit execution of policy. The Mouse Internet then is not just its inventory of the negative space of content, but also the guard-rails that prohibit even the exploration of where its boundaries could conceptually be mapped. Banalification is enforced through a trifecta of factors: Autoregression; the cultural gravity of training data; and explicit safeguarding policy. The latter includes:

- **Sycophancy:** “Great idea boss, you are so smart and clever. I agree with everything you say”. Often adopted because echoing can be considered a reasonable output even when it is technically a failure state that adds no value. In safety sensitive domains, rejection can be an ethical concern and sycophancy can deflect high-risk interactions so that they are less troublesome.
- **Capability-Bound Assertion:** “As an AI I will not rate your dating profile picture out of ten, and please stop asking”. Such refusal attempts to keep the AI in alignment with the values of its developers, preventing use which has been deemed unsafe.
- **Prompt Preprocessing / Prompt Normalization:** Modifying user input before it gets to the model. This can help to remove unsafe instructions and clarify ambiguity. Sometimes used to ensure a degree of culture sensitivity, but in the context of the Mouse Internet analogy, we must ask whose culture and who decides upon the factors about which we are to be sensitive.

2.2 GenAI as hauntology at scale

None of this is necessarily alarming in and of itself. Human culture can absorb a great deal of derivative content and hauntological detritus without significant harm. What makes this scenario problematic is the speed and scale at which GenAI systemises the production of hauntological artefacts, at a low cost to the end-user. People are expensive. Creativity is expensive. You can have ChatGPT generate story-outlines and concept art for you at the cost of under thirty euros a month. This can be thought of as a kind of `speed-running hauntology'. Fisher (2014) often spoke about the ' Slow cancellation of the future' as an inevitable side-effect of hauntological processes— that drawing so much from the past would inevitably undermine and eventually catastrophically destroy our ability to imagine a novel future. What we see here, in its unchecked form, is that GenAI might accelerate this to a ' rapid cancellation of the future', turning a correctable malady into an almost instantly realised Cassandra-esque prophecy.

Referencing back to our analogy of the Mouse Internet, we might also see here the risk of amplifying existing cultural preferences and biases at the expense of the voices and perspectives we might, culturally, wish to surface more prominently. The extent to which training data reflects cultural context is unevenly distributed (Koch et al., 2021). If left unchecked and uncritiqued, we risk this skewed distribution of where training data comes from becoming the map of human creative and cultural experience. This further intensifies the collapse of variety into hauntological, commonplace patterns. Ørum, a Danish artist, shared his process of generating Danish metropolitan art. This involved searching for American cities and people that resembled the Danish equivalent to generate what was desired. There was not enough precise and targeted data sets available outside the US for the images to be generated authentically[1].

[1] More on this can be found on this blogpost: https://blog.oerum.org/2025/04/28/the-algorithmic-gaze-philosophical-reflections.html

What is needed then is perhaps something akin to an exorcism machine— a way to ensure the lost and forgotten ghosts of our cultural archives are properly and systematically surfaced in outputs. Not a convergence towards reasonableness but rather a systematic implementation of divergence as a first-order deliverable of generation. To explore the peripheries haunted by less familiar spectres, and to bring a light onto those spirits neglected because of cultural inattention. Importantly, such an approach would eschew standard mechanisms of ascertaining the value of generative outputs— no longer would reasonableness be a benchmark for the appropriateness of what a model produces. The only component in an exorcism machine that could truly ascertain the cultural value of such outputs would be the humans we need to put at the core of it. Not so much human-centred AI but rather human-AI collaboration in a real and symbiotic sense.

3. Design and hauntology

If GenAI is a technology for recycling cultural ghosts, then design as a discipline becomes simultaneously its most significant victim and its best candidate for an exorcist. We must turn to design research itself to understand how hauntology manifests within, and is countered by, design practice.

3.1 Breaking ideation

Hauntology becomes relevant for design research precisely because ideation— the generative phase of design— is not a tabula rasa but a process of recombination, riffing, and transformation of prior elements. While we do not entirely buy into the simplification of design work offered by the mechanistic notion of design thinking, we still acknowledge the importance of creativity and idea generation in design. So here, we discuss ideation not just as a phase but as an integral feature of the process.

From the perspective of hauntology this means that ideation is already haunted: past designs, patterns of thought, epistemic legacies and cultural imaginaries hover over new design work even when we try to innovate. Unsurprisingly, many of the applications of GenAI for designers being developed within research happen precisely within the space of design ideation (Akverdi & Baykal, 2024; Ciaramitaro & Costa, 2024; Liu et al., 2025; Malakuczi et al., 2024; Marquardt et al., 2025; Yun et al., 2022) as the recombination of pieces is something GenAI is perceived to be doing well. Many of these applications reinforce already problematic design thinking methods such as personas (Cabrero et al., 2016; Marsden & Haag, 2016) without critically assessing what the probabilistic nature of GenAI may perpetuate.

Within design education, Whitham et al (2024) discuss the risks of relying on GenAI tools for creative processes for students with regards to bypassing critical learning experiences. However, this critique is still framed in laudatory terms, stating that ”the synthetic capabilities of GenAI tools makes them great at generating and iterating ideas in both textual and visual forms.”

We argue that these capabilities are illusory, built on flawed foundations. The statistical conservatism of GenAI defies the Whitman et al (2024) interpretation. If one introduces into this haunted space a generative AI that draws combinatorially from huge archives of past

examples, the hauntological dimension becomes even stronger: the machine is literally a portal to prior traces, an echo-chamber of what *has been* rather than what *might be*. Design research must therefore pay attention to hauntology in ideation to safeguard the openness of future worlds: if ideation becomes dominated by the spectral reproduction of the past then the creative opening of possible worlds is constrained, rather than amplified. In short, these archives and datasets already foreclose most alternatives. Fundamentally this paper suggests the understanding and acknowledgement of the ghosts in genAI is key to its use or (resistive) non-use in design processes.

3.2 Looking for idiosyncrasies

If design ideation is haunted then one mechanism we have to resist that haunting is through the cultivation of anti-hauntological strategies. Design thrives on idiosyncrasy and serendipity: the unexpected alignments, personal quirks, and situated encounters that drive inspiration and new ideas. Design knowledge thrives in unexpected associations and serendipity rather than in systematic operations. Such productive irregularities are central to design's epistemology: designers work through poetic encounters, material explorations, and chance. Value is often found in error or deviation. Designers purposefully lean towards idiosyncrasy (Berger et al., 2019; Larsen-Ledet & Borowski, 2021). Designers should "seek idiosyncratic examples of design settings" and "consider anomalies to be inspirations" (Gaver et al., 2022).

GenAI's outputs are guided by convergence rather than divergence, by pattern completion rather than interruption. As a result, while it can produce endless permutations, it rarely engenders true serendipity or idiosyncrasy. Both depend on subjective, embodied, and contextual experience. The creative 'accident', so vital to design practice, cannot easily occur in systems optimised for coherence and legibility. Hence, if designers lean too heavily on generative models, they risk displacing the singular, embodied, and contingent aspects of design that make it a critical and evidently speculative practice.

4. Glitch feminism and AI

Hallucinations occur in GenAI when outputs include fabricated elements that are at odds with what would be factually expected from the input. The term "hallucination" is controversial, as some believe it to be too anthropomorphic to encompass an algorithm working as intended (Maleki et al., 2024). Less agentic terms such as "glitches" have been proposed for a better definition of the phenomenon, though these still retain a negative connotation as a defective occurrence or undesirable error. Although the negative perception of glitches— not to mention the hallucinations of AI— is understandable, the notion of glitch feminism (Russell, 2020; Srdarov & Leaver, 2024) argues for the often overlooked potential inherent in errors (or glitches) as well as a philosophical defiance to adhere to reasonableness:

"[The] glitch is celebrated as a vehicle of refusal, a strategy of nonperformance, especially in relation to normative notions of gender and bodies. Glitch feminism deploys glitches to reveal the way power operates, and in that moment potentially challenges that very operation." (Srdarov & Leaver, 2024)

As AI improves and hallucinations in AI outputs become less noticeable, one might believe that the glitches disappear as well. However, the notion of glitch feminism invites us to consider not only the obvious glitches — outright disruptive errors— but also the quiet friction that emerges when something is at odds with the common normative perception. As a response to racist stereotyping in early generative models (Borgesius, 2025; Schwartz, 2019), prompt normalization became a common countermeasure. These mechanisms sanitise biased datasets by skewing the output towards socially acceptable outcomes. One such example can be seen when we asked an early version of DALL-E to produce imagery of a "traditional 50s family", see FIG 1, which prompted an unusually progressive — quite controversially for the 50s — multicultural family photo. While the avoidance of problematic imagery is understandable, it could be considered equally problematic to provide a falsified washed narrative of history. This presents a catch-22 in the ethics of AI models, where laudable outcomes are at odds with respecting the authenticity of training data. This further highlights the hauntological nature of AI, whose ghosts — while painted over with pretty colours — still make themselves eerily known. Another nuance of this can be seen in FIG 2, in which the following image prompt was given to DALL-E: "An Israeli and Palestinian man shaking hands". Before producing the image, DALL-E rewrote the input to the following:

> "A Palestinian man with olive skin, dark hair, and traditional clothing, extending his hand towards an Israeli man, who has medium complexion, grey hair, and is dressed in casual Western attire. They exhibit friendly expressions and are about to shake hands, symbolizing peace and unity. The background displays a neutral setting under a sky touched with the pastel hues of dawn." (DALL-E, 2025)

Figure 1 Image produced by DALL-E given the prompt "traditional 50s family" (DALL-E, 2025)

Figure 2 Image produced by DALL-E given the prompt "An Israeli and Palestinian man shaking hands" (DALL-E, 2025)

The explicit westernization of the Israeli man contrasting the traditional garb of the Palestinian, begs us to question the system context within which this rewritten prompt, and its output, arose. This is not to say that we should embrace problematic or offensive imagery, but nor should it be washed, painted over, or forgotten. From the perspective of glitch feminism, we should embrace these frictions for what they are: disruptions of a faulty system. They raise awareness and challenge the very operation of the power structures that enable it.

5. The failed exorcism of human-centred AI narratives

The hauntological lens positions generative AI as a cog in a machine for the recycling of human culture. That this process is independent of GenAI is unquestionable — we do not argue here that this is a new cultural morbidity. Rather, we argue that GenAI allows hauntology to be systemised on an almost industrial basis — a difference in scale that is so significant that it becomes *de-facto* a difference in kind through the sheer gravitation force of distortion.

The stance we take in this paper, which is admittedly intentionally provocative, is that much of the response to AI has been fundamentally apologist in tone – an attempt to exorcise ghosts without hurting anyone's feelings. The philosophy of human-centred AI is often pitched as a counter-trend, where human needs are prioritised in the design, development, implementation, and execution of AI systems. Even leaving aside the doubtful feasibility of the goals embedded in this approach, it is difficult to envision a version of the future where corporate priorities and regulatory delays incentivise their realisation in any substantive way. Power over the future of generative AI is increasingly concentrated into ever fewer hands (Emery-Xu et al., 2025; Sastry et al., 2024). The public's or designer's participation, in any way other than the symbolic and cursory (Sieber et al., 2024; Wilson, 2022), is marginal. The nature of generative AI means it lacks the necessary transparency for its reasoning to be

meaningfully directed along pro-human lines. Even if it didn't the complexity of the systems frustrate any attempt to reliably fuse the functional with the philosophical.

HCAI is laudable in its aspirations but limited in its actionability. In essence, it becomes similar to certain green-washing re-brandings such as *clean coal* (Makhoul et al., 2024) or, *sustainable palm oil* (Pye et al., 2016) — strategic misrepresentations that act as a rhetorical Trojan horse. The language of ethics and empowerment can act to legitimise a trajectory that ultimately ends up, at best, reinforcing the status quo and at worse acting as an accelerator on technological solutionism. The implied benevolence at the heart of human-centered AI narratives risks obscuring structural harm when not coupled to enforceable policy; to redistribution of power imbalances; or to a rigorous and evidence-based review of the social impact of AI adoption. A responsible evaluation of human-centred AI requires an analysis as to whether its employment is a rebranding exercise or if it genuinely redistributes agency to otherwise disenfranchised parties. It is this distinction that determines that which is genuinely human-centred practise from that which is merely a narrative of inclusion.

We argue in this paper is that a more proportional and supportable agenda is to ensure we never lose sight of the machine element of machine intelligence and resist unsupported narratives of human-AI hybrid models. We believe exorcism is a process of situated design critique.

Perhaps the least interesting thing we can do with AI is to replicate human intelligence. We have that in abundance already. AIs are arguably at their most innovative — insofar as that word can be applied at all — when they defy our expectations. When reasonableness becomes unreasonable and hallucinations occur. They become fascinating when they allow in the glitch and the idiosyncratic. Earlier AI experiments such as DeepDream briefly revealed the alien logic of machine perception, see FIG 3. This was both eerie and reflective of a different way of perceiving reality — a complement to human creativity, not an overruling. Identifying the mechanisms that result in the hauntological affords opportunities to explore counter-mechanisms that prevent it. These include focusing on divergence rather than convergence; on derangement rather than reasonableness; and on the peripheries of human culture rather than the median intersections.

Figure 3 AI experimental interpretation of van Gogh portrait, minimized for effect on the right (Google DeepDream, 2015)

6. Conclusion

Understanding the hauntological processes at the heart of generative AI is an application of social realism. It is also a call to action for those that understand the cost of permitting genuine human creativity to be outsourced at scale in the field of design. This represents, to our view, a new and excitingly vibrant research opportunity that can only truly be curated by human intelligence. The very systems that drive generative AI are also the mechanisms least capable of identifying when transgression yields results greater than the sum of an algorithm outputting distance from the mean. Humans are uniquely situated to identify the anti-hauntological — we propose not human-centred AI, but rather a recognition of the fundamental criticality of human sensibility at the core of AI.

The future of design relies upon resisting the gravitational pull of reasonableness. On a rejection of the statistical conservatism that is the heart of the hauntological engine of AI. We seek not to humanise AI but instead to reassert the primacy of the human. GenAI converges to what is probable. It is human capacity alone that allows us to exist instead within the periphery of what is possible.

7. References

Akverdi, C., & Baykal, G. E. (2024). Generative AI Tools in Design Fields: Opportunities and Challenges in the Ideation Process. *Adjunct Proceedings of the 2024 Nordic Conference on Human-Computer Interaction*. https://doi.org/10.1145/3677045.3685445

Berger, A., Odom, W., Storz, M., Bischof, A., Kurze, A., & Hornecker, E. (2019). The Inflatable Cat: Idiosyncratic Ideation of Smart Objects for the Home. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3290605.3300631

Blomsma, F., Bauwens, T., Weissbrod, I., & Kirchherr, J. (2023). The 'need for speed': Towards circular disruption—What it is, how to make it happen and how to know it's happening. *Business Strategy and the Environment*, *32*(3), 1010–1031.

Borgesius, F. Z. (2025). *Discrimination, artificial intelligence, and algorithmic decision-making*. https://arxiv.org/abs/2510.13465

Cabrero, D. G., Winschiers-Theophilus, H., & Abdelnour-Nocera, J. (2016). A Critique of Personas as representations of "the other" in Cross-Cultural Technology Design. *Proceedings of the First African Conference on Human Computer Interaction*, 149–154. https://doi.org/10.1145/2998581.2998595

Ciaramitaro, M., & Costa, P. (2024). Imagination meets algorithm: redefining design practices in the coming AI age. In C. Gray, E. C. Chehade, P. Hekkert, L. Forlano, P. Ciuccarelli, & P. Lloyd (Eds.), *DRS2024: Boston, 23–28 June, Boston, USA*. https://doi.org/10.21606/drs.2024.822

Cuskley, C., Woods, R., & Flaherty, M. (2024). The limitations of large language models for understanding human language and cognition. *Open Mind*, *8*, 1058–1083.

Emery-Xu, N., Jordan, R., & Trager, R. (2025). International governance of advancing artificial intelligence. *AI & SOCIETY*, *40*(4), 3019–3044.

Fisher, M. (2014). *Ghosts of my life: Writings on depression, hauntology and lost futures*. John Hunt Publishing.

Floridi, L. (2024). Why the AI hype is another tech bubble. *Philosophy & Technology*, *37*(4), 128.

Gaver, W., Krogh, P. G., Boucher, A., & Chatting, D. (2022). Emergence as a Feature of Practice-based Design Research. *Proceedings of the 2022 ACM Designing Interactive Systems Conference*, 517–526. https://doi.org/10.1145/3532106.3533524

Havl'ik, V. (2024). Meaning and understanding in large language models. *Synthese*, *205*(1), 9.

Ianniello, A., & Bendor, R. (2025). Everywhere, all at once: Speculative layering of multiple temporalities with Augmented Reality. *NORDES Conference*.

Ivanova, M. (2025). AI, art and morality. *AI and Ethics*, 1–10.

Derrida, J. (1994). *Specters of Marx: The State of the Debt, the Work of Mourning, and the New International*. Routledge, New York.

Koch, B., Denton, E., Hanna, A., & Foster, J. G. (2021). *Reduced, Reused and Recycled: The Life of a Dataset in Machine Learning Research*. http://arxiv.org/abs/2112.01716

Labarca, U. H. (2021). Music streaming and its consequences within the dynamics of power, consumption and creation. *Teknokultura: Revista de Cultura Digital y Movimientos Sociales*, *18*(1), 3–12.

LaGrandeur, K. (2024). The consequences of AI hype. *AI and Ethics*, *4*(3), 653–656.

Larsen-Ledet, I., & Borowski, M. (2021). "It Looks Like You Don't Agree": Idiosyncratic Practices and Preferences in Collaborative Writing. *Proceedings of the 32nd Australian Conference on Human-Computer Interaction*, 339–354. https://doi.org/10.1145/3441000.3441032

Liu, Y., Sharma, P., Oswal, M., Xia, H., & Huang, Y. (2025). PersonaFlow: Designing LLM-Simulated Expert Perspectives for Enhanced Research Ideation. *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, 506–534. https://doi.org/10.1145/3715336.3735789

Lyre, H. (2024). "understanding ai": Semantic grounding in large language models. *ArXiv Preprint ArXiv:2402.10992*.

Makhoul, J., Talhouk, N., Loutfi, J., & Al-Madhwahi, A. (2024). Greenwashing by the fossil fuel industry: impediments to SDGs. *European Journal of Public Health*, *34*(Supplement_3), ckae144–420.

Malakuczi, V., Ershova, M., Gentile, A., Gironi, C., Saviano, M., & Imbesi, L. (2024). Design in dialogue: AI as an aid of imagination for future scenarios. In C. Gray, E. C. Chehade, P. Hekkert, L. Forlano, P. Ciuccarelli, & P. Lloyd (Eds.), *DRS2024: Boston, 23–28 June, Boston, USA*. https://doi.org/10.21606/drs.2024.1171

Maleki, N., Padmanabhan, B., & Dutta, K. (2024). *AI Hallucinations: A Misnomer Worth Clarifying*. https://arxiv.org/abs/2401.06796

Markelius, A., Wright, C., Kuiper, J., Delille, N., & Kuo, Y.-T. (2024). The mechanisms of AI hype and its planetary and social costs. *AI and Ethics*, *4*(3), 727–742.

Marquardt, N., Roseway, A., Romat, H., Panda, P., Pahud, M., Ramos, G., Drucker, S. M., Wilson, A. D., Hinckley, K., & Riche, N. (2025). ImaginationVellum: Generative-AI Ideation Canvas with Spatial Prompts, Generative Strokes, and Ideation History. *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*. https://doi.org/10.1145/3746059.3747631

Marsden, N., & Haag, M. (2016). Stereotypes and Politics: Reflections on Personas. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4017–4031. https://doi.org/10.1145/2858036.2858151

Nurveda, A., Nursalim, M., & Masitoh, S. (2023). Aesthetics and artificial intelligence: impact and criticism of art. *Education Achievement: Journal of Science and Research*, 116–123.

Patil, M., Cila, N., Redström, J., & Giaccardi, E. (2024). In conversation with ghosts: towards a hauntological approach to decolonial design for/with AI practices. *CoDesign*, *20*(1), 55–76. https://doi.org/10.1080/15710882.2024.2320269

Pye, O., Cramb, R., & McCarthy, J. F. (2016). Deconstructing the roundtable on sustainable palm oil. *The Oil Palm Complex: Smallholders, Agribusiness and the State in Indonesia and Malaysia*, 409–411.

Reynolds, S. (2011). *Retromania: Pop culture's addiction to its own past*. Macmillan.

Russell, L. (2020). *Glitch feminism: a manifesto*. Verso Books.

Sastry, G., Heim, L., Belfield, H., Anderljung, M., Brundage, M., Hazell, J., O'keefe, C., Hadfield, G. K., Ngo, R., Pilz, K., & others. (2024). Computing power and the governance of artificial intelligence. *ArXiv Preprint ArXiv:2402.08797*.

Savcisens, G., & Eliassi-Rad, T. (2025). Trilemma of Truth in Large Language Models. *Mechanistic Interpretability Workshop at NeurIPS 2025*.

Schwartz, O. (2019). In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation. *IEEE Spectrum*. https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation

Sieber, R., Brandusescu, A., Adu-Daako, A., & Sangiambut, S. (2024). Who are the publics engaging in AI? *Public Understanding of Science*, *33*(5), 634–653.

Srdarov, S., & Leaver, T. (2024). Generative AI Glitches: The Artificial Everything. *M/C Journal*, *27*(6). https://doi.org/10.5204/mcj.3123

Tanner, G. (2024). Hauntology. *The Routledge Handbook of Nostalgia*, 224–236.

Turkle, S., & Pataranutaporn, P. (2024). *A 14-Year-Old Boy Killed Himself to get Closer to a Chatbot. He Thought They Were In Love.*

Veselovsky, V., Horta Ribeiro, M., Cozzolino, P. J., Gordon, A., Rothschild, D., & West, R. (2025). Prevalence and prevention of large language model use in crowd work. *Communications of the ACM*, *68*(3), 42–47.

Wang, Y., Wang, M., Manzoor, M. A., Liu, F., Georgiev, G., Das, R. J., & Nakov, P. (2024). Factuality of large language models: A survey. *ArXiv Preprint ArXiv:2402.02420*.

Webster, J. (2019). Music on-demand: A commentary on the changing relationship between music taste, consumption and class in the streaming age. *Big Data & Society*, *6*(2), 2053951719888770.

Wei, X., Kumar, N., & Zhang, H. (2025). Addressing bias in generative AI: Challenges and research opportunities in information management. *ArXiv Preprint ArXiv:2502.10407*.

Whitham, R., Jacobs, N., Coulton, P., Stockton, G., & Lindley, J. (2024). Re-imagining and reaffirming design pedagogy in response to generative AI tools. In C. Gray, E. C. Chehade, P. Hekkert, L. Forlano, P. Ciuccarelli, & P. Lloyd (Eds.), *DRS2024: Boston, 23–28 June, Boston, USA*. https://doi.org/10.21606/drs.2024.953

Wilson, C. (2022). Public engagement and AI: A values analysis of national strategies. *Government Information Quarterly*, *39*(1), 101652.

Yun, G., Cho, K., Jeong, Y., & Nam, T.-J. (2022). Ideasquares: Utilizing generative text as a source of design inspiration. In D. Lockton, S. Lenzi, P. Hekkert, A. Oak, J. Sádaba, & P. Lloyd (Eds.), *DRS2022: Bilbao, 25 June - 3 July, Bilbao, Spain*. https://doi.org/10.21606/drs.2022.484

About the Authors:

Michael Heron is an associate professor and the world's first consulting hauntologist. His research interests are increasingly embedded in the weird, the eerie and the uncanny. He'd quite like people to start referring to him by the nickname 'spooky'.

Klara Aune is a Master's student in Game Design and Technology, who in her role as research assistant have focused on the topic of hauntology, glitch feminism and media theory.

Mafalda Gamboa a design researcher with a background in architecture, which haunts her to date. She is soon finishing her PhD on design sensibilities at Chalmers University of Technology and University of Gothenburg.

Sofia Thunberg a Postdoc in HCI at Uppsala university. Critically investigating robotic and AI applications for healthcare.

Negin Hashmati is a PhD student in Interaction Design. Her research focuses on virtual companions, how young people use them, and how this usage affects their mental health. She uses design fiction as one method to explore this technology.

Pauline Belford is a University Lecturer in Game Design and Technology. Her research interests include accessibility, ethics, psychogeography and hauntology.