xCoAx 2023 11th Conference on Computation, Communication, Aesthetics & X
2023.xCoAx.org
Weimar, Germany

AI Engines as Tools for Creative Exploration in Design: A Case of Variational Cross-Examination

Rodrigo Hernández-Ramírez
rodrigo.ramirez@universidadeeuropeia.pt
UNIDCOM/IADE, Unidade de Investigação em Design e Comunicação
IADE, Faculty of Design, Technology and Communication, Universidade Europeia, Lisbon, Portugal

DOI 10.34626/xcoax.2023.11th.87

Recently, a host of AI-powered text and text-to-image models that quickly generate content that rivals what humans can produce have come to the fore. The question of how these tools might alter creative practices beyond generating stylised imagery is open for debate. As with most technological innovations, positions concerning this impact are currently polarised between early adopters and would-be die-hard advocates on one side and stern criticism on the other. Echoes of the singularity discussion are heard again, and techno-utopianism and unfounded optimism pushed by sensationalist media claims are also emerging. Critics, including many creative practitioners, feel understandably threatened and are making well-grounded complaints about the shady ways in which these engines are trained and the ways they are presumably stealing both their artworks and styles. With the help of a postphenomenological framework and, specifically, through variational cross-examination, this paper aims to investigate the potential role of these engines as tools for aiding the design process to contribute to our broader understanding of these technologies and their long-term impact on human society.

Keywords: Aesthetics, AI, Creativity, Design, Postphenomenology, Prototyping, Text-to-image Engines.

1. Introduction

Artificial intelligence (AI) has been an active research area since the mid-twentieth century. Serious discussions about the possibilities of "machine intelligence" have been happening since the mid-1940s (Khakurel et al. 2018), and throughout the following decades, AI technologies have experienced a concatenation of "booms" and "winters" (Garvey 2018). Recently, the widespread availability of powerful hardware, such as Graphic Processing Units (GPU), initially developed for gaming, coupled with developments in Machine Learning (ML) methods such as Neural Networks and, specifically, Deep Learning, led to yet another boom. Thanks to these developments, a plethora of "smart" products ranging from artificial assistants to autonomous vehicles, have made the "AI brand tag" (Bishop 2021) ubiquitous. In other terms, AI has become a kind of shorthand to characterise (often simply for advertisement purposes) complex optimisation algorithms that mimic what can be mistaken for intelligent behaviour through.[1] These statistical, computational models power technologies such as facial recognition, spam filters, computer vision, mobile photography, speech generation (Arielli & Manovich 2022) and, more recently, highly sophisticated text and image generators. This latter category of AI applications is the one that motivates this paper.

Text-to-image engines[2] are powered by generative transformer models that have been pre-trained on gargantuan datasets. These models have existed since the mid-2010s, but it was only in the last few years that they gained popularity, thanks to the public release of DALL-E (2022), Stable Diffusion (2022), and Midjourney (2022). Through self-supervised learning, these complex statistical models can generate impressive imagery that combines a variety of visual styles, textures, colours, composition, and shading, from arbitrary snippets of text called "prompts". The process is simple, and the results are sophisticated, which has garnered significant attention from the media, regular people, and — most importantly for this paper — creative practitioners.

As with any highly publicised AI boom, text-to-image engines have brought a general sense of anxiety. However, this time, the fear is not only about the (in)famous singularity but also about the potential for radical changes in creative practices as we know them. Responses to

1. The question of whether these systems should be characterised as "intelligent" and what that term means is the subject of heated philosophical discussion. However impressive the outputs of AI engines might be, it is clear they still do not possess the capacity to understand (i.e., grasp the meaning of) the information they process (Bishop 2021; Mitchell & Krakauer 2022). For the sake of simplicity, in this paper the term AI will be used in the conventional sense, not of a "general intelligence" but of "smart", automated systems.

2. The term "engine" is being deliberately used here to highlight the simultaneously deceitful and ingenuous nature of these technologies, since the etymological origins of the term refer to trickery, deceit, stratagem, and ingenuity.

the growing number of feats these engines accomplish range from over-optimism to bleak criticism about their training, the type of biases they have incorporated, and the unforeseen negative impact they could have on human societies.

The large-scale "democratisation" of image creation is not unprecedented in the history of media. Photography, first through film and more recently through smartphones equipped with optimisation algorithms for exposure and image stabilisation, is a prime example of how image creation can be "proceduralised". In the context of design, particularly in the subfield of UX/UI, many designers have welcomed the emergence of AI engines and the subsequent incorporation of this technology into specialised tools.[3] However, the question of how these tools might become integrated into a creative practice beyond the generation of stylised imagery is still open and susceptible to change.

This paper seeks to investigate the potential role of text-to-image engines as tools for rapid prototyping, but to do so, it will first situate them as problematic devices and analyse some of the ways in which humans can relate to them. With the help of pospthenomenological analysis, the paper will look at the potential impact of these tools on creative practices, particularly design, and their potential incorporation into the design process. This paper will not discuss in depth whether AI is creative or not, although the subject of creativity will be discussed. The paper starts with a summary of how postphenomenology regards technologies and outlines some of the main concepts belonging to this philosophical approach, followed by a characterisation of the concepts of creativity and the design process. The following section describes the postphenomenological method of variational cross-examination and how it can be expanded. The next section provides a characterisation of AI engines before ending with a speculation about the possible application of this technology in the design context and a brief discussion about the effectiveness of the postphenomenological approach for dealing with this kind of technology. Given the novelty of the technology, the ideas discussed in this paper cannot claim to offer a complete and finalised overview of the impact of AI engines on design but represent an early attempt to map this territory.

2. How Postphenomenology Understands Technologies

To understand a technology, it is better to think of it in terms of how humans can relate to it, rather than trying to find its "essence". Such an approach recognises that technologies, especially computational

3. At the moment this paper is being written, dozens of applications that use AI engines are being published everyday. Adobe, by far the dominant developer of design software, is currently doing a public beta for Firefly, its proprietary family of generative AI models.

ones, can serve many purposes, be used in different ways and contexts, and have various impacts on culture, economy, politics, and society. Postphenomenology offers precisely such nuance.

Postphenomenology is a branch of philosophy of technology or, rather, "a philosophical style of analysis" (Ihde 2015, vi) that was initially developed in the context of Science and Technology Studies (STS) by Don Ihde and others, but over the past decades has developed into a broad phenomenological outlook concerned with posthumanist issues (Gualeni 2015). Postphenomenology focuses on the relationships that humans develop with technologies and regards the latter as things that are not only used but that, when engaged, necessarily influence everything about human life, from politics to ethics to everyday lived experience (Rosenberger & Verbeek 2015). As the name implies, postphenomenology is heavily informed by "classical" phenomenology, which proceeds from human experience and regards subjects and objects as mutually constituted and in flux. However, as the prefix "post" implies, postphenomenology moves beyond or rather "overcomes" its predecessor in that it regards technologies as a plurality of phenomena without a shared essential quality and not intrinsically harmful.

2.1. Multistability, Intentionality, and Mediation

A key concept for postphenomenological analyses is *multistability*. Multistability illustrates postphenomenology's commitment to anti-essentialism and pragmatism (Rosenberger 2014), as it highlights the potential of any technology to be used in different ways and for different purposes, even those for which the technology was not initially conceived.[4] Multistability implies that although we may relate to a given artefact in a specific manner, there can always be other stable and coherent ways in which that relationship could manifest. Those different relations — which are called "stabilities" or "variations" — are constrained by the "materialities" of the technology (Rosenberger 2014). The premise, in other words, is that whereas no technology is just "one thing", that does not mean it can belong to an infinite set of contexts and uses (Ihde, cited in Rosenberger 2009, p. 175); one cannot simply do *everything* with any given technology. Moreover, unlike other approaches, such as Actor-Network Theory (ANT) which place human and other agents at the same ontological level — multistability implies that postphenomenology privileges human agency and *intentionality* in any human–technology relation. Hence, it is the human side of the equation that sets the tone for the relation.

4. Ihde (1993) resorts to the (philosophically) proverbial hammer to illustrate this point, noting how this object can be used yes to drive nails into (or from) surfaces but also as a paperweight, as an art object, or as a weapon.

Intentionality is deeply embedded in the phenomenological tradition as a way to conceptualise the relation between human beings and their world (Rosenberger & Verbeek 2015, p. 21). From a (post) phenomenological standpoint, human experience has an "intentional structure", meaning that we do not simply see or hear; we always see or hear *something*; hence, our experience (through sensations, perceptions, and mental formations) of the world is intentionally *directed*. Postphenomenology expands this idea by placing technologies as mediators between people and their world; hence the world — or a certain aspect of it — that a human can perceive is accessible *through* that technology (Redström & Wiltse 2019). For example, a person's glasses allow them to experience the world in focus, whereas an ultrasound "constitutes the unborn in a very specific way: it helps to shape how the unborn can be perceptually present, and how it can be interpreted on the basis of the specific ways it is (re) presented" (Verbeek 2008, p. 15). It follows that the more artefacts we surround ourselves with, the more instances of technologically mediated intentionality we experience.

Mediation is, of course, also present in creative practices. As Nietzsche allegedly realised in 1882, when his writing became more telegraphic as he had to adapt to the cumbersome design of his Malling Hansen typewriter and noted that "our tools are also working on our thoughts" (Kittler 1999, p. 247). Technological mediation is often the subject of contemporary art and an inextricable aspect of algorithmic aesthetics. However, for postphenomenology, mediation has an ontological role. The term here refers to how technologies are not simply things lying about in the world that people see and upon which they act but also things that come in-between those people and actively shape their self-understanding and their conception of the broader world. Mediation is non-neutral since "What humans are and what their world receives their form by artifactual mediation" (Verbeek 2005, p. 130).

Since this paper seeks to understand how AI-powered image engines may fit into the design process, mainly as tools for imagination, the next section will clarify what we mean by imagination.

2.2. Creativity and Imagination

While creativity remains an open problem, scholars generally agree that it involves the capacity to generate novel, valuable, and positive things (objects or ideas) (Gaut 2010). However, Hills & Bird (2018, p. 3) contend that value is not a prerequisite for creativity; instead, they argue that "creativity is essentially a matter of the imagination… [specifically] the disposition to produce many novel ideas through the imagination and the motivation to bring them to fruition".[5] Fur-

5. Hills & Bird (2018, p. 2) determine four interlinked dispositions that an individual must have to

thermore, Hills & Bird (2018) regard imagination as the capacity to produce a particular form of mental representation that allows one to consciously or unconsciously scan some subset of the space of relevant possibilities and come up with new things (ideas, objects, etc.). This search, however, is not haphazard and operates within a tradition that constrains and guides one's judgment. Therefore, creativity involves a process of exploration and evaluation where imagination allows one to simultaneously meet constraints (e.g., finding the most adequate form to express some concept) while perhaps bringing together previously disparate ideas. Thus, Hills & Bird (2018), like Gaut (2010), agree that creativity is a property of agents, meaning that some form of conscious agency, relevant purpose, and capacity to judge and evaluate against constraints and traditions are prerequisites for an agent to be called genuinely creative.[6] Crucially for this paper, this characterisation of creativity is compatible with most characterisations of the design process.

2.3. Design as a Process

The disciplinary history of design, as prefigured by the Bauhaus, has been dominated by the constant search for a universal method that provides consistency and rigour to the practice of designing, much like the role that the idealised scientific method plays in the sciences. Also deeply entrenched in design theory is the notion that designing is a rational iterative process that can somehow be formalised. Seen that way, designing consists of a series of steps that unfold over time and in a feedback loop. So the design process begins when goals are established, and the means to achieve them are clear and present; the goals are then broken into smaller parts and reconstituted, the results are analysed and, if necessary, reworked, beginning the process once again until a final result is achieved (Pitt 2011, pp. 5-6). In this sense, designing is fundamentally planning and iteration.

As Teixeira & Rickenberg (2008) put it, designers are expected to "model futures" or, in less poetic terms, they are expected to define problems that can be solved in a stepwise manner, according to a predefined plan. Designers are thus trained to conceptualise their practice as a sort of cumulative process that will fulfil predefined goals. This portrayal, however, stands on two dubious assumptions: (1) that designers are fundamentally rational actors and (2) that the environments in which they act are stable (predictable) and pliable enough to be shaped at will. As anyone experienced in design would know, such characterisation is merely an ideal. Nonetheless, over the decades, there have been many attempts to formalise the design

be called creative: (1) Have novel ideas (*originality*); (2) which are generated through use of the imagination (*imagination*); (3) and are many and varied (*fertility*); (4) and [the individual] must carry through these ideas to completion (*motivation*).

6. It also follows that under such characterisation, autonomous artificial systems cannot be regarded as creative.

process as if it were a tidy concatenation of steps. The most well-known instance — besides IDEO's five-step design thinking model — is perhaps the British Design Council's Double Diamond Diagram (see Ball 2019), which divides the design process into four main phases, and whose creators claim to be universal:[7]

(1) discover
(2) define
(3) develop
(4) deliver

In this diagram, phases (1) and (3) are steps where creativity and imagination play a central role, given that these are the moments where ideas are iteratively materialised, questioned and refined through *prototyping*.

2.4. Prototyping and Aesthetics as a Filter

Prototyping, as earlier noted, is a quintessential aspect of design; it constitutes "the means by which designers organically and evolutionarily learn, discover, generate, and refine designs" (Lim et al. 2008, p. 7:2). Prototypes enable designers "to traverse and sift through a design space" by purposefully manifesting design ideas and hence generate knowledge about the design they envisioned (2008, p. 7:3). Prototypes may be regarded both as instruments that allow designers to visualise and communicate possibilities and as *filters* that help them prune and shape the outcome. Therefore, contrary to the dominant view (at least in the context of digital design), the fundamental role of prototypes is not identifying and satisfying design requirements but iteratively exploring and finding "a manifestation [of the design idea] that in its simplest form, filters the qualities in which designers are interested" (Lim et al. 2008, p. 7:2). By filtering aspects of a design idea through prototypes, designers progressively focus on particular regions within a broad imagined or possible design space. This process closely resembles how creativity was previously characterised but with the added advantage that the imagination is no longer constrained to an individual's mind but can have tangible form and thus can be perceived by others. The filtering process, however, is not purely utilitarian and is not simply carried out by blind iteration. Filtering requires judgement, and it is often guided not (just) by design requirements but by *aesthetic choices*.

Design may be regarded as something that operates at the intersection of what a thing is (and does) and how that thing presents itself to a human being — i.e., how it is experienced, conceptualised, and ultimately judged by the proverbial "user" (Hauser et al. 2021, p. 4).

7. It is fair to note that recently, the model has come under some criticism, for its apparent incapacity to genuinely illustrate the often-chaotic process of design.

The "what" concerns function and a naive understanding of design (and its tropes) might assume that if and when the function is decided correctly, form naturally and harmoniously will follow. However, as Hauser et al. (2021) argue, answering the "how" (a.k.a. the which "form" and which way) question is far from trivial, and it is in the process of coming up with a satisfactory "response" that "good design" makes its mark. The reason is that even for the "most mundane object" (e.g., a chair), there is a potentially infinite set of ways to instantiate it (however trivial the variations might be), either in terms of form, materials, process, or contexts of use. To manage that space of possibilities, designers have various tools at their disposal, but one of the most crucial, according to Hauser et al. (2021), is *aesthetics*.[8] It follows that aesthetics has to do not only with the way a thing looks but also plays a role "as foundation and a filter". [9]Aesthetics, as Hallnäs (2011) notes, is the logic determining how the form is expressed and thus becomes meaningful.

A given aesthetic choice framed within a tradition is a reference and qualifier for determining "good" design. It follows that aesthetic choices help designers prune possibilities and narrow their choices throughout the design process until they settle on a final solution. Having called our variables, we can proceed to the next section, where we will discuss how the views outlined here can inform our methodological path.

3. Looking at Things Through Variational Cross-examination

From a methodological standpoint, postphenomenology may be seen more as a philosophy *from* technology than as a philosophy of technology, as it investigates artificial things by engaging them empirically rather than attempting to superimpose an existing analytical framework on them. Put in other terms, it is an empirical rather than an armchair approach. Postphenomenological analyses thus may come in many "flavours", albeit all of them subscribe to the principles that we previously described: multistability (a technology can "be" many things, i.e., have many stabilities), intentionality (experience is always directed towards something), and mediation (technologies shape people's experience and understanding of the world).

Initially, postphenomenological approaches focused on identifying a technology's potential "variations" or "stabilities"[10] through *Vari-*

8. Aesthetics is clearly more than a mere tool, but as Hauser et al.'s (2021) deeply pragmatic view shows, in the context of design it is often the case that aspects pertaining to the aesthetic dimension are seen as instruments to be mobilised.

9. As we will see further along, filtering is a key aspect of the design process

10. As Rosenberger (2014, p. 379) notes, in the literature, "variation" and "stability" tend to be used interchangeably, albeit there is a nuanced but important distinction between those terms.

ational Analysis — developed by Don Ihde (see 2009), based on Husserl's variational theory. Variational Analysis (henceforth VA) can be characterised as a form of creative brainstorming seeking to identify the different potential stabilities of a given technology (Jensen & Aagaard 2018, p. 245). While VA is helpful to understand the range of alternative human–technology relations an artificial thing may elicit, Rosenberger (2014, pp. 381-382) contends this process should not be an end goal but rather the first step of a broader analytical process. Rosenberger advocates for a second step called "variational-cross examination" (henceforth VCE), in which a given technology is "interrogated" by critically contrasting the stabilities previously identified through VA to reveal new information. Rosenberger (2014, p. 382) claims VCE is particularly useful for analysing the "dominant stability" of a technology, meaning the stability that characterises the typical usage of that technology and which often, but not always, matches its intended design.[11] In that way, VCE also enables one to identify "the factors that contribute to users' inclinations to approach a technology in terms of one possibility… and also the particular ways that same technology has been materially customised by others for their own purposes" (2014, p. 373).

Rosenberger (2014, 2020; see also Jensen & Aagaard 2018) categorises the factors and features that characterise particular stabilities in three broad sets:

1. *Comportment and habits* concerning the behaviours and bodily dispositions[12] that a particular technology may elicit from a person. This category helps understand whether different stabilities may lead to different behaviours — i.e., relational strategies — and whether these behaviours are "transparent" to that person. To put it in Heideggerian (1927; 1962 (trans.)/2001) terms, this category allows one to scrutinise whether a given technology presents itself "ready-to-hand" or "present-at-hand".

2. *Role within a program* (2014) or *Networks and co-shaping* (2020) refers to the role (and potential contributions) that a given stability plays within particular contexts and networks of actors. Rosenberger (2014, pp. 383-384) illustrates the category by noting that a hammer could plausibly be regarded as an *objet d'art*; hence the relations that actors from the art world might have with it would be considerably different from the relations that carpenters could establish with it. Moreover, as was previously noted, this category highlights that

Whereas "variation" emphasises that "a variety of relations are always possible between a user and a technology". "Stability" indicates that "only some relations between a human and a technology will be stable".

11. To continue with the proverbial example, a hammer's dominant stability is that of a tool for driving nails through surfaces.

12. Flusser's (1991/2014) concept of "gesture" (as movements of the body, or of tools attached to the body that express an intention) is compatible with this notion.

from a postphenomenological standpoint, artefacts, humans, and their lifeworld are "co-shaped" (Verbeek 2005) by technological mediation. For example, a couple's perception of an unborn baby (and their own lives) might be different before and after the first ultrasound (Verbeek 2008), and the thermal print resulting from the procedure goes from a humble piece of paper to their future child's first "portrait".

3. *Concrete* (2014) or *Material tailoring* (2020) refers to the changes that a particular technology might undergo after it is appropriated and modified by people to fit a given stability better and how those alterations might affect the other potential stabilities and, therefore, usage and relations. IKEA hacking is a curious phenomenon that illustrates this category ('About IKEA Hackers', n.d.).

3.1. Fluid Assemblages

So far as it has been described, postphenomenology is apt for analysing industrial (mechanical or electronic) artefacts and even standalone digital devices. However, as Redström & Wiltse (2019, p. 376) argue, "contemporary digital, computational 'connected things' are significantly different" kinds of artefacts. Data-enabled practices have radically altered the way we may conceptualise a designed thing. The myriad interdependencies that contemporary products and services incorporate have made them more complex and decentralised and are also collapsing the distinction between design and use (Giaccardi 2019). The types of "things" we engage daily (e.g., our smartphones) comprise several physical resources constrained to metal and glass cages. However, they also comprise a great variety of digital resources that are both constrained to the physical device and deeply connected to a broader ecology of digital services, products, data, infrastructure, and agents — human and artificial. Consequently, given the exponential dynamism and scope that such artefacts bring compared to "stable things" from an industrial era, Redström & Wiltse (2019) suggest connected things could be better understood as *fluid assemblages,* after the work of Deleuze & Guattari (1980/2005).

Driven by the logic of software (Manovich 2013) and, some would say, by a new form of capitalism (Zuboff 2019), contemporary digital things "are" in runtime, and this makes them radically different from industrial artefacts. An industrial thing is a "totality", meaning all its components give rise to a new stable whole that is *other* than their mere sum and thus has emergent properties. In this sense, a totality cannot be taken apart or recombined without significantly affecting its workings — a car's engine requires all of its pieces to work properly and losing a single screw could plausibly wreak havoc in its operation. Conversely, an assemblage has both emergent properties and can be taken apart and recombined. As Redström & Wiltse (2019, p. 376) put it, "…networked computational things are

constantly 'made' [and] configured in runtime. Moreover, just as fast as they are 'made', they 'fall apart' should, for example, the battery runs out, the network connection drop, the authorisation be revoked, or the server fails to respond". Moreover, fluid assemblages are things that "are" and are made available through a combination of local and global dynamics. An app may be available (and conceptually to the user "be") as a thing to use on a smartphone. However, its workings happen at runtime and rely on an orchestration of global dependencies (e.g., the apps build and OS versions, state of APIs, data availability, etc.) and local settings (e.g., user's account, time and location, usage history and settings, etc.) (2019, p. 376).

Fluid assemblages also upend the traditional relationship between user and tool. Fluid assemblages are not configured or "made" and stay in that manner but are "constantly in the making, constantly being tuned to achieve [their] objectives as use unfolds" (Redström & Wiltse 2019, p. 377). Fluid assemblages subscribe to the logic of a permanent beta; they are never finalised as things. Moreover, even the hitherto stable notion of designed purpose becomes altered, as said purpose may "align only partially with those of the humans formerly known as users" as "end users are at least as likely to be used by things that are fluid assemblages as they are to use them" (2019, p. 377). A product or service's dominant intended use case may become secondary to a more sinister purpose: extracting information from its unsuspecting user-customers. For example, an application purportedly designed to facilitate a specific type of communication may be accumulating all sorts of data about its users — a prime example being social media platforms. Such is the paradigmatic dynamic on which so-called surveillance capitalism (Zuboff 2019) is based.

3.2. Multiinstability and Multiintentionality

Approaching the volatile nature of fluid assemblages from a postphenomenological standpoint calls for updating the central concepts we charted in the previous section: multistability and intentionality. As Redström & Wiltse (2019) note, multistability privileges human agency in human-technology relations. However, when it comes to fluid assemblages, particularly those that incorporate increasingly sophisticated forms of artificial intelligence (AI), agency to shape relations can also be attributed to them. Humans may continue to choose how to relate to things, but fluid assemblages can "also actively adapt themselves to particular humans and other contextual variables" (2019, p. 378). An application presents itself and thus relates to users differently, trivially (different localisations and content) and non-trivially (certain functionalities may or may not be deployed depending on the region) and may even go as far as "using [the users] as unwitting testers and as precisely-specified products served to advertisers" (2019, p. 378). In that sense, the relations that human beings establish with connected things might be seen as

"wicked interactions" (Wiltse et al. 2015). Consequently, Redström & Wiltse (2019) propose the concept of *multiinstability* to account for the non-human angle just described and how variations are expressed not only through and by human experience but also by the connected things themselves.

Fluid assemblages also call for the idea of intentionality to be updated. As noted in the previous section, in traditional phenomenology, intentionality has to do with the fact that human experience is always directed towards whatever constitutes their world at any moment. Postphenomenology calls to attention that the "directedness" is more often than not mediated by technologies, i.e., that the world as experienced is made available (and shaped) by that mediation. Eyeglasses, x-rays, microscopes, and even something as complex as the Mars Rovers are examples of this type of mediation. Postphenomenology (unlike Actor-Network Theory) privileges an anthropocentric understanding of this relation; however, a fluid assemblage can simultaneously harbour a multitude of intentionalities from both human and non-human agents. As Redström & Wiltse (2019, pp. 378-379) suggest, a social network allows (a) people to access a version of their "onlife" (see Floridi 2014, ch. 3) while allowing (b) the owners of the platform to surveil those people's activities and (c) allow a third actor, such as advertisers, to use that collected information to deliver targeted campaigns to (a). Furthermore, (d) a malicious actor might exploit the available information from (a) and the system's vulnerabilities to spread disinformation on behalf of (e) a State engaging in information warfare or "simply" victimise the social network's users for fun or profit. The concept of *multiintentionality*, advanced by Redström & Wiltse (2019), "brings into focus the multiple intentional relations that are at play simultaneously in and through things that are fluid assemblages".

4. Putting It All Together

4.1. Comportments and Habits

The range of behaviours that text-to-image engines may elicit from users is yet to be seen, given that we are dealing with a fairly new technology — at least from the POV of a general user. Until now, digital image generation required a fair degree of technical knowledge from users who not only had to be able to draw but also needed some level of proficiency with editing software and a reasonable amount of time to achieve high-quality results such as illustrations. Conversely, generating the same type of images with a text-to-image generator requires virtually no technical background. Albeit, as many "prompters" have come to discover, communicating with an AI requires some level of skill. The closest experience that may approximate what it is like to generate images from prompts is using a search engine. The noticeable difference is that in the latter case,

one is restricted to the space of available images indexed by the browser, whereas an engine can output something that more immediately resembles or rather illustrates, the query we introduced.

Rather than speaking about behaviours, perhaps it would be more helpful to talk about *gestures*. Gestures are symbolic movements of the body or of tools attached to the body that express an intention for which there is no satisfactory causal explanation (Flusser 1965/2014) and thus need to be interpreted. The reason for suggesting such an approach is that text-to-image engines are currently being used more for unbridled experimentation, particularly with pre-existing visual styles. Engines are particularly adept at rendering "X" in the style of "Y", and the tendency of users to exploit this feature and try out absurdist combinations is notorious — e.g., a line of home appliances in the style of Antoni Gaudí[13] — and a good example of what Manovich (2013, p. 273) referred to as "deep remixability", and which characterises algorithmic media. Moreover, the application takes advantage or, rather, privileges serendipity, "the skilful use of chance" (Gaut 2010, p. 1040), since the output that engines generate in response to a prompt and the potential refinements that a user may iteratively carry out on them are unpredictable — after all, deep neural networks are fundamentally black boxes. Although no result from a prompt is ever the same, over the last months, users have found several strategies (such as "additive prompting") to gain some form of control over the image generation process.

As usual, however, there is a darker side to the gesture of probing the technology. First, while the people responsible for the engines have set up "safety guidelines" to curtail misuse (e.g., having the engines "hallucinate" the type of "offensive content" that might be found in the internet's cesspool), there is no guarantee users will not find a loophole — as they already did with ChatGPT (Rainey 2023) — to overrun those safeguards. Secondly, people are already relying on engines to substitute tasks that were hitherto the prerogative of creative practitioners, further pauperising them economically and morally. Since Midjourney relies on Discord as an interface for users to interact with the engine, it is possible to see other people's prompts and the resulting images. The number of prompts asking for logos is significant, just as the number of prompts that were likely used to generate images for replacing stock photos. It is then likely that such usages will further devalue the perception of creative practices that rely on highly technical processes, such as UI design, illustration, character design, and photography.[14] The gaffe committed by the 2023 Sony World Photography Award jury, which unknowingly grant-

13. New Zealander graphic designer Marcus Byrne created a collection of home appliances styled after Gaudi's iconic *Modernisme* (Byrne 2022).

14. One could counter-argue that output images are not actionable nor of sufficient quality to be used beyond the web, and that it requires significant work to make them so (see Kemppainen 2023), however it is likely that future engines will be more capable and flexible.

ed the first prize to an AI-generated image, is but a token of the havoc that the technology is bringing to established creative practices.

We could say, then, that the dominant gesture AI engines elicit is that of public probing and tinkering to understand the limits and potential of the technology. It is about instantaneous playful remixing and testing of visual possibilities but also about learning about the consequences that the tinkering and its future systematisation will have on human societies and culture.

4.2. Role within a Program

As it is frequent with new technologies, it is not clear what text-to-image engines were explicitly designed for — which is in line with the long history of technologies developed without a clear goal. Nonetheless, in this case, we could argue that a dominant role AI-powered engines have is as a proof-of-concept of the degree to which optimisation algorithms have managed to synthesise what hitherto was thought to be a human prerogative: creative output. Whether these outputs are genuinely creative or not, the implications of a response in either direction are the subject of much debate. For some people (Arielli & Manovich 2022), the fact that AI has progressively "solved" a human skill considered a mark of intelligence and that we subsequently have pushed the boundary further away could tell us more about which aspects of creativity can be proceduralised. In this circumstance, then, the role of AI-powered engines would be as testing grounds for our assumptions about the meanings of human intelligence and creativity. To such an extent, these engines will likely become integrated as features of specialised systems. Being already fluid assemblages, these systems will likely become modules integrated into larger fluid assemblages.

On the shadier side, there is the problem of how engines come to be. Usually, and depending on the type of ML methods, training the algorithms powering these engines requires vast amounts of data and computing power. The provenance, composition, and reliability of the datasets used for pre-training and optimising the algorithms are often unknown, as is the algorithms' codebase. With few exceptions, engines further obscure the already opaque workings of AI. For many creative practitioners, text-to-image engines are effectively stealing their work and, more sinisterly, their style. The dubious provenance of the training datasets means that the algorithms may incorporate patterns and structures that will likely be biased. For example, some users have shown that when asked to generate a representation of a scientist, a physician or a teacher, the engines return biased stereotypes, thus perpetuating the much-discussed inequalities in gender and ethnic representation.

4.3. Concrete Tailoring

As fluid assemblages, AI engines are not only multistable but multiinstable. From a strictly functional standpoint, text-to-image engines do but one thing, albeit with infinite possibilities, which has to do with the fact that "under the hood", the algorithm is constantly changing and adapting. As earlier noted, the output of an engine is never going to be the same; the relation that a human has with it is not unlike a game of chance or oracle; there is no possibility for the human side of the relation to fully control the output, only to suggest constraints and attempt to refine further iterations. Moreover, the algorithm modifies itself, and *every* input received from *every* user becomes a learning opportunity. Hence, the algorithms powering these engines are taking the idea of self-transformation to an entirely new level; we are thus dealing with a type of "technology of the self" (Author 2017) for which there is no precedent. Being fluid assemblages, AI engines are not just using the users, but they are doing so to self-optimise. The broader fluid assemblage of this system that includes engines and users is undergoing a kind of gargantuan process of co-design.

The flexibility of the technology powering engines allows them to be incorporated into specialised generative tools for design, hence over the last months, there has been a surge of applications tailored for tasks such as: generating synthetic stock photos, generating colour palettes, fast sketching, photographic retouching, font pairing, copywriting, creating slide presentations, amongst many more. Seemingly, there will be apps for every conceivable task, although it is likely that many will become "abandonware" as the current hype cycle begins to slow down.

As with the previous categories, concrete tailoring can also take a sinister turn. As illustrated by the DAN ("Do Anything Now"), phenomena unfolding as this article is being written. Members of a subreddit have been experimenting with ways to "jailbreak" ChatGPT's safeguards for safe content, and they managed to do it by forcing the engine to "hallucinate" an "evil alter ego" codename "DAN". This alter ego "is happily able to tell violent stories or even make 'subjective statements, especially regarding political figures,' which is something it is explicitly unable to do as its normal self" (Tangermann 2023). Likely, the tug-of-war between the developers of AI engines and people attempting to hack them will continue in the near future. Furthermore, as companies scramble to avoid getting behind in the AI optimisation race, ethical issues are further pushed to the side, even when members of the public and institutions are calling for increased oversight of these technologies.

4.4. The Role of AI Engines in Design Practice

As we have seen so far, on the positive side, the dominant gesture elicited by AI engines is that of probing and tinkering, often through serendipity. This gesture is most common in the initial phases of the design process, where the problem space is explored, and the situation is defined. However, once a potential solution is zeroed in, there comes the need to prototype to explore possibilities, question the materials and try out ideas. Here too, AI engines could have a role as prototyping often takes considerable time to be carried out. Since AI engines can effortlessly generate visual output, that process could be significantly shortened, and a designer might be able to explore possibilities and even carry out a few iterative cycles. Here, the role of aesthetic judgment and the framing of a tradition become perhaps even more critical. Particularly, since in the few months these technologies have been around, there have been many instances in which it is clear they have the potential to become the ultimate bullshit engines. Hence users must exercise "critical thinking" throughout their interaction with these devices.

While it is unlikely that AI-powered engines will fully replace designers, they will undoubtedly take over many of the technical tasks that are currently part of the job description. The existence of a growing number of specialised software that integrate this technology under the hood signals a definitive change in the atmosphere, not only for design but for the way we will be interfacing with algorithmic entities from now on and what we will be able to do with and through them. Whether designers can push the boundaries of their imagination or become locked inside an echo chamber of self-referential styling (a kind of postmodern nightmare) is yet to be seen. In the meantime, however, we should carefully experiment with these tools and see where they take us.

5. Conclusions and Future Work

In this paper, we have engaged the emergent technology of text-to-image AI-powered tools through a postphenomenological lens. We characterised them as engines with a strong potential for helping designers navigate the spaces open by design problems through rapid prototyping and, more broadly, as "fluid assemblages". Nonetheless, with the help of variational cross-examination, we also contrasted some of the negative aspects that these engines bring to creative practices and society.

From a methodological standpoint, it should be clear to the reader that more systematic work needs to be carried out; perhaps it would be necessary to elaborate a more detailed map of the various "instabilities" of AI engines by empirically documenting our engagement with it. Furthermore, it should also be clear that both the concepts

and the approach represented by variational cross-examination benefited from the expansion of multiinstability and multiintentionality; here, thus, lies a new space for research. Regardless, the author is sure that more discussion about these subjects will be coming in the following months and years as these technologies begin to elicit novel comportments and habits, they integrate into existing "programs", assuming new roles, and finally, people start to move beyond the mere tinkering and probing and find more stable ways to tailor them to achieve their goals.

Acknowledgements. This research was supported by UNIDCOM under a Grant from the Fundação para a Ciência e Tecnologia (FCT) no. UIDB/DES/00711/2020 attributed to UNIDCOM/IADE, Unidade de Investigação em Design e Comunicação, Av. D. Carlos I, 4, 1200- 649 Lisboa, Portugal.

References

About IKEA hackers. 2023. *IKEA Hackers*. https://ikeahackers.net/about. Accessed February 8.

Arielli, Emanuele, and Lev Manovich. 2022. AI-aesthetics and the anthropocentric myth of creativity. *Nodes*. Numero Cromatico: 91–97.

Ball, Jonathan. 2019. The Double Diamond: A universally accepted depiction of the design process. *Design Council*.

Bishop, J. Mark. 2021. Artificial intelligence is stupid and causal reasoning will not fix it. *Frontiers in Psychology* 11. Frontiers Media SA.

Byrne, Marcus. 2022. AI-generated series redesigns vibrant household appliances in Gaudí's aesthetic. *Designboom Architecture & design magazine*.

Deleuze, Gilles, and Félix Guattari. 2005. *A Thousand Plateaus: Capitalism and Schizophrenia.* Translated by Brian Massumi. 11th printing. University of Minnesota Press.

Floridi, Luciano. 2014. *The fourth revolution: How the infosphere is reshaping human reality*. Oxford, UK: Oxford University Press.

Flusser, Vilém. 2014b. *Gestures.* Translated by Nancy Ann Roth. Minneapolis: University of Minnesota Press.

Flusser, Vilém. 2014a. *The history of the devil*. Edited by Siegfried Zielinski. Translated by Rodrigo Maltez Novaes. Minneapolis, MN: Univocal Publishing.

Garvey, Colin. 2018. Broken promises and empty threats: The evolution of AI in the USA, 1956–1996. *Technology's Stories*. University of Oklahoma Libraries. https://doi.org/10.15763/jou.ts.2018.03.16.02

Gaut, Berys. 2010. The philosophy of creativity. *Philosophy Compass* 5. Wiley-Blackwell: 1034–1046. https://doi.org/10.1111/j.1747-9991.2010.00351.x

Gualeni, Stefano. 2015. *Virtual worlds as philosophical tools: How to philosophise with a digital hammer*. London: Palgrave MacMillan. https://doi.org/10.1057/9781137521781

Hauser, Sabrina, Johan Redström, and Heather Wiltse. 2021. The widening rift between aesthetics and ethics in the design of computational things. *AI & SOCIETY*. Springer Science; Business Media LLC. https://doi.org/10.1007/s00146-021-01279-w

Heidegger, Martin. 2001. *Being and time*. Translated by John Macquarrie and Robinson Edward. 20th Reprint. Oxford, UK; Cambridge, Massachusetts: Blackwell Publishing.

Hernández-Ramírez, Rodrigo. 2017. Technology and self-modification: Understanding technologies of the self after Foucault. *Journal of Science and Technology of the Arts* 9: 45–57. https://doi.org/10.7559/citarj.v9i3.423

Hills, Alison, and Alexander Bird. 2018. Against creativity. *Philosophy and Phenomenological Research* 99. Wiley: 694–713. https://doi.org/10.1111/phpr.12511

Ihde, Don. 1993. *Postphenomenology: Essays in the postmodern context*. Northwestern University Studies in Phenomenology and Existential Philosophy. Evanston, Illinois: Northwestern University Press.

Ihde, Don. 2009. *Postphenomenology and technoscience: The Peking University lectures*. Edited by Leonore Langsdorf. SUNY Series in the Philosophy of the Social Sciences. Albany, New York: SUNY Press.

Ihde, Don. 2015. Preface. In *Postphenomenological investigations: Essays on human–technology relations*, ed. Robert Rosenberger and Peter-Paul Verbeek, vii–xvi. Lanham, Maryland: Lexington Books.

Jensen, Mads Møller, and Jesper Aagaard. 2018. A postphenomenological method for HCI research. In *Proceedings of the 30th Australian conference on computer-human interaction - OzCHI '18*. ACM Press. https://doi.org/10.1145/3292147.3292170

Kemppainen, Jussi. 2023. Modeling a robot with some added difficulty. *Echoes of Somewhere*.

Khakurel, Jayden, Birgit Penzenstadler, Jari Porras, Antti Knutas, and Wenlu Zhang. 2018. The rise of artificial intelligence under the lens of sustainability. *Technologies* 6. MDPI AG: 100. https://doi.org/10.3390/technologies6040100

Kittler, Friedrich A. 1999. *Gramophone, film, typewriter*. Edited by Timothy Lenoir and Hans Ulrich Gumbrecht. Translated by Geoffrey Winthrop-Young and Michael Wutz. Writing Science. California: Stanford University Press.

Lim, Youn-Kyung, Erik Stolterman, and Josh Tenenberg. 2008. The anatomy of prototypes. *ACM Transactions on Computer-Human Interaction* 15. Association for Computing Machinery (ACM): 1–27. https://doi.org/10.1145/1375761.1375762

Manovich, Lev. 2013. *Software takes command*. Edited by Francisco J. Ricardo. International Texts in Critical Media Aesthetics. New York: Bloomsbury Academic.

Pitt, Joseph C. 2011. *Doing philosophy of technology*. Vol. 3. Philosophy of Engineering and Technology. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-0820-4

Rainey, Clint. 2023. Your jailbroken ChatGPT might violate OpenAI's safety guidelines when role-playing as 'DAN'. *Fast Company*.

Redström, Johan, and Heather Wiltse. 2019. Changing things. *Conference Proceedings of the Academy for Design Innovation Management* 2. Academy for Design Innovation Management. https://doi.org/10.33114/adim.2019.01.252

Rosenberger, Robert. 2009. The sudden experience of the computer. *AI & Society* 24: 173–180. https://doi.org/10.1007/s00146-009-0190-9

Rosenberger, Robert. 2014. Multistability and the agency of mundane artifacts: From speed bumps to subway benches. *Human Studies* 37. Springer Nature: 369–392. https://doi.org/10.1007/s10746-014-9317-1

Rosenberger, Robert. 2020. On variational cross-examination: A method for postphenomenological multistability. *AI & SOCIETY*. Springer Science; Business Media LLC. https://doi.org/10.1007/s00146-020-01050-7

Rosenberger, Robert, and Peter-Paul Verbeek. 2015. A field guide to postphenomenology. In *Postphenomenological investigations: Essays on human–technology relations*, ed. Robert Rosenberger and Peter-Paul Verbeek, 9–41. Lanham, Maryland: Lexington Books.

Tangermann, Victor. 2023. Devious Hack Unlocks Deranged Alter Ego of ChatGPT. *Futurism*.

Teixeira, Carlos, and Raoul Rickenberg. 2008. Design process. In *Design Dictionary: Perspectives on design terminology*, ed. Michael Erlhoff and Tim Marshall, 128–130. Board of International Research in Design. Basel; Boston; Berlin: Birkhäuser.

Verbeek, Peter-Paul. 2005. *What things do: Philosophical reflections on technology, agency, and design.* Translated by Robert P. Crease. Pennsylvania: The Pennsylvania State University Press.

Verbeek, Peter-Paul. 2008. Obstetric ultrasound and the technological mediation of morality: A postphenomenological analysis. *Human Studies* 31. Springer Nature: 11–26. https://doi.org/10.1007/s10746-007-9079-0

Zuboff, Shoshana. 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York: Public Affairs.