BALANCING BETWEEN SCIENCE AND ART: GENERATIVE AI IN ARCHITECTURE

Konstantinos Michos
Aristotle University of Thessaloniki
konmichos@gmail.com

Abstract

Artificial intelligence has recently entered every area of culture, revealing untapped potential as well as fears that technology may advance in unpredictable ways. In generative imagery, AI tools allow images tailored to the user's wishes to be created almost instantly from a single prompt. While DALL-E, Midjourney and Stable Diffusion have brought AI image generation into popular culture, related algorithms have been employed in very different fields: in scientific research (the M87* and Sgr A* black hole images, though generation did not include the totality of the image) and in commercial applications, including advertising campaigns by the Coca-Cola Company, BMW and Nutella. The first case challenges our beliefs about the abil-

Digital Age in Semiotics & Communication, Vol. IX, 2026, 99–122
Under CC BY 4.0
https://doi.org/10.33919/dasc.26.9.5

ity of AI to produce new scientific knowledge; the second unsettles our ideas about art and creativity.

The versatility of AI allows it to be used in very different contexts. One type of image, however, blurs the line between science and art: architectural design. With many commercial tools already available, AI-generated architecture is hard to classify. By studying a set of AI images of imaginary Santorini resorts that became popular in Greece, this article attempts to shed light on the way such images are read. The images are then tested against a set of semiotic, cultural and epistemological validation criteria in order to separate their artistic qualities from the information they convey about reality. The results show that examining such images from multiple points of view provides safer readings of AI-generated content and can reduce opportunities for misuse.

Keywords: AI-generated images, architecture, sensorial trance, cognitive surrender, semiotic–cultural–epistemological validation

1. Introduction

Artificial intelligence has lately turned out to be a science-fiction dream come true. From convincing conversations to extraordinary images, commercially available AI applications challenge our established beliefs and call for a reorganization of the way we think, work and produce. Moreover, as hardware and software both evolve rapidly, it is still unclear when progress will begin to slow down and reach a plateau that would allow us to reflect on its achievements.

Or is this really so? While some researchers praise the new tool, a significant number point out various problems arising from its use, including sustainability (Lakshmi et al. 2024): AI is resource-intensive, and scaling it up depends on strict control of efficiency. But even if progress can somehow be sustained, the greatest concern is arguably raised by its ethical implications: intellectual property rights, shifts in the job market and the gradual loss of human skills.

This study does not attempt to address every such aspect, but focuses on one: the epistemological status of AI-generated images. It was the first generation of AI image models – DALL-E (2021), Midjourney (2022) and Stable Diffusion (2022) – that drew attention to the epistemic status of such images. Within a short time, however, more and even more advanced models

appeared: ChatGPT (which already excelled as a language model), Google's Nano Banana, Qwen, Flux and others, with various forks multiplying the possibilities further. These are merely the models best known in popular culture because they are commercially available; they do not include proprietary algorithms designed for more specialized environments, such as Adobe's AI tools for image editing or models for scientific purposes (Event Horizon Telescope Collaboration 2019), where algorithm CHIRP was employed to fill the visual parts missing after combining available astronomical observations.

While early criticism focused on the misuse of the technology to create fake news or compromising images of celebrities and others, it soon became apparent that a tool so powerful at passing off the fake as real could challenge many more areas in which reliable knowledge of the truth is important. Meanwhile, its ability to construct realistic-looking scenes of evidently impossible situations sparked enthusiasm about the new creative frontiers that could be reached. AI can easily construct compositions like Philippe Halsman's *Dalí Atomicus* (1948), which shows Salvador Dalí with three flying cats – although that was an actual photograph, not an artificial image. Unfortunately, the training of the models usually involved the unauthorized use of intellectual property, which led many to condemn their use in all artistic endeavours. Whether AI can be truly creative has also been questioned.

As AI becomes more popular and more widely experimented with, its status – whether as a tool for reaching scientific truth or for creating imaginary art – is being challenged. More interesting still, there is at least one area of applied knowledge that relies on both these properties: architecture.

Architectural designs need to be highly creative in order to satisfy specific needs or particular wishes. On the other hand, an extremely innovative design that is impossible to build, physically or financially, is of no use, except perhaps as an experiment. AI tools quickly became popular among architectural firms and designers for their fast and innovative solutions to classic problems. The renowned Austrian architecture firm Coop Himmelb(l)au, for example, has its own AI division, Deep Himmelb(l)au, which is considered one of the most vivid examples of the use of AI in architecture (Chaillou 2025). Such tools are even more popular among people with limited access to expensive resources; as a result, many social media accounts of freelance architects focus on AI-generated designs.

The research question of this article is whether AI-generated images can be perceived as reliable representations of reality or simply as icons of the imaginary. Architecture was chosen as a field of application precisely because of its hybrid nature, which requires both scientific and artistic

expertise. Their combination produces artefacts that are more difficult to check against conventional criteria (for example, for AI hallucinations). To this end, a set of images of architectural designs on the island of Santorini that gained some online popularity in Greece was selected for analysis. A combination of semiotic, cultural and epistemological criteria is proposed, whose joint fulfilment more accurately reflects a healthy relation between AI images and reality. The results show that context is crucial in such cases if the public is to read the images properly.

2. The current status of AI-generated images

The text-to-image models mentioned above offer a wide range of advantages for image creation: quick results, no need for special skills and increasing accuracy. These qualities are often emphasized in optimistic narratives, especially with regard to creativity and democratization. Bainss (2023: n.p.) claims that "[b]y embracing the potential of AI in pop culture while being mindful of its limitations and drawbacks, we can work toward a future where technology and culture are more inclusive, representative, and meaningful for all". Generative AI is also less stigmatized in the natural sciences, mainly because of its problem-solving capacity (Joksimovic et al. 2023).

Naturally, there are also less optimistic views. Apart from the ever-relevant issues of intellectual property and personal data (Smits & Borghuis 2022; Chesterman 2024), and of deepfakes and fake news (Helmus 2022; Matthews & Kidd 2023; Romero Moreno 2024), the rhetoric surrounding AI has also been discussed. Leone (2024: 430–431) observes that scientific fields and technologies are often perceived as "augmented" through the use of certain signs, such as a doctor's lab coat. Such signs build a narrative around the technology; in the case of AI, the rhetoric is based on industrial secrecy, that is, on "black boxes" (432–433). Along similar lines, Russo, Schliesser, and Wagemans (2024) argue that a shift to a glass-box approach, in which systems are fully transparent and inspectable, would greatly improve the explainability of AI – a move that is, however, intrinsically linked to the internalization of ethics within the epistemology of AI.

Durán and Jongsma (2021) examine the role of trust in the use of AI in medicine. Like Leone and Russo, Schliesser, and Wagemans, they identify opacity as the core of the problem, but they sidestep the demand for transparency. Since complete transparency is difficult or even unattainable, they propose computational reliabilism as an alternative, which relies not on complete transparency but on other indicators of reliability. In their view,

the most important of these is expert knowledge: AI results should be accompanied by an expert reading. Human critical thinking is thus not excluded from the process but incorporated into it in an elevated epistemic function.

Anichini and Kotras (2024) also study the conditions under which AI produces epistemically sound results, concluding that scientific objectivity is largely shaped by local constraints, that is, by context, and is therefore "not a given property but an emergent feature" (Anichini & Kotras 2024: 631). Expert knowledge clearly belongs to this context. Scotti (2024) is also concerned with trust in AI images but takes a different approach: instead of treating images as passive items detached from their context (while acknowledging that this context exists), she proposes that AI images are both instituting and instituted. To assess the status of AI images fully, we should therefore focus not only on the way they are created but also on how they affect their use, which in turn shapes the system that produces them.

Another interesting view is presented by Alvarado (2023), who explicitly defines AI as an epistemic technology. An epistemic technology does not contain the truth; it is simply a tool for inquiring into, manipulating and operating on epistemic content, that is, data. In this way he sidesteps the epistemological risks involved, but he also reduces AI to a tool that functions solely according to the user's input, leaving aside the generation of novel knowledge observed by others.

Others have expressed their concerns more openly. Campolo and Crawford (2020) note that the comparison between AI and alchemy goes back as far as 1965 and stress the need to move from an "enchanted determinism" to "disenchantment" (Campolo & Crawford 2020: 5). Stuart Russell (in Ford 2018: 42) maintains that the deep learning hypothesis about how AI works is still a kind of magic. Yampolskiy (2020) argues that even if AI could somehow explain some of its decisions, these explanations would not necessarily be comprehensible to people.

It appears, then, that the opacity of AI fuels the discussion about the epistemology of AI images. It is also apparent, however, that traditional approaches may in any case be unable to describe what is real and true within an AI framework. The next section discusses different definitions of the "real" in order to shed light on the new challenges brought by the age of AI.

3. Towards a new definition of truth

The comedian and television host Stephen Colbert coined the term *truthiness* on his show in 2005 to describe the growing use of emotional rhetoric,

coupled with unchecked facts, in US politics. The word was named Word of the Year by the American Dialect Society and by Merriam-Webster. At the time, AI as we know it today did not yet exist. Indeed, the discussion then centred on the way narrative interacts with our willingness to accept facts, especially when they clash with our beliefs, rather than on the epistemology of AI.

Some of the points raised in that discussion nonetheless apply to AI today. First, the results of generative AI depend heavily on context. They therefore cannot be presented as absolute truth, but rather as a *version* of the truth which, when accompanied by sufficient context and a glass-box view of the algorithmic procedures, can at least restore our trust in the mechanism that produced it. In other words, truth in AI is sensitive and multifaceted. Second, AI emerged at a time when the world was already interconnected through the dense digital web of the internet. Consequently, every AI accomplishment and artefact sends lightning-fast ripples through the digital world. When readers are exposed to systems this influential, they are more inclined to assume that what they are shown is the truth and later to incorporate it into their everyday conversations. This self-reinforcing cycle turns AI-generated knowledge into the norm and feeds it back into the training of future models.

This is not to suggest that readers bear no responsibility. Even if facts do not form the most compelling narrative, consuming AI-generated content requires – like any other activity – an informed approach. It cannot be stressed enough that, in the case of AI, such informed reading defines the limits of its epistemic function.

Shin (2026) elegantly articulates the need for a new definition. He argues that AI systems have moved beyond being mere tools for detecting misinformation to actively reshaping the epistemological foundations of truth itself, and he introduces the concept of *algorithmic truth* to describe this quality, that is, "the systematized and probabilistic processes through which AI systems render claims as factual" (Shin 2026: 1553). He draws a clear distinction between the *verification* and the *interpretation* of truth: the former relies on "probabilistic reasoning, pattern recognition, and anomaly detection" (Shin 2026: 1553), carried out by AI through inference, while the latter involves a deeper understanding of knowledge. AI systems, he observes, excel at the former but often struggle with the latter (Shin 2026: 1554).

This effect is greatly magnified by the omnipresence of AI tools and the digital distribution of their content, which leads to a paradigm shift away from traditional epistemologies (Shin 2026: 1553). While Shin does not

deny the efficiency and scale of this shift, he worries about "narrowing epistemic authority and silencing alternative knowledge systems" (Shin 2026: 1555). He also identifies the epistemic opacity of AI systems as a problem (Shin 2026: 1556) and agrees that AI findings are context-sensitive: "Ground truth datasets are often framed as epistemically stable, yet they are socially constructed" (Shin 2026: 1556). This is a problem not only from an epistemological but also from a political point of view, since, although "AI systems are often framed as neutral, they are embedded in sociotechnical contexts that reflect and reproduce longstanding inequalities and power asymmetries" (Shin 2026: 1557). Such asymmetries undermine democratic institutions, which also suffer from a lack of trust in the mechanisms of generative AI.

Alongside algorithmic truth, Shin describes *algorithmic trust*, which involves people trusting systems they do not understand. It is presented as an evolutionary step in the epistemology of AI: for the sake of efficiency, modern systems of knowledge production cannot afford delays, but if we are to avoid entrusting ourselves to black boxes, this algorithmic trust should rest on inclusive, pluralistic infrastructures of verification that do not reproduce systemic asymmetries "in whose knowledge counts" (Shin 2026: 1558).

The lack of a suitable definition of truth in AI-driven digital environments is inextricably linked to the lack of tools for identifying it. Language models are getting close to passing the Turing test, but telling a real image from an artificially generated one requires different skills. Our definition of the real depends on our ability to track it, and to do so we must completely disengage the real from the photorealistic.

4. Photorealistic vs real

Photography has been described as a process strongly linked to reality (Bazin 1967; Sontag [1977] 2005; Walton 1984). Early theorists identified this property because the first photographic devices were strictly deterministic. As technology progressed, this link was severed, and today we are losing the ability to distinguish the real from the photorealistic. Caporusso, Zhang, and Carlson (2020) report eye-tracking experiments in which people correctly identified only 60% of AI-generated images – a percentage that has probably fallen further since, because the study, though recent, predates the explosive development of generative image models.

AI-generated images come into being in a very different way from light inscribed on a photosensitive surface. The process at work in most current models can be summarized as follows: an initial dataset of images is chosen, in which the images have been paired with content labels (by human

annotators or through AI classification); noise is added to the images, in a process similar to physical diffusion (hence the name *diffusion models*); the noise is then removed in a reverse process to produce new images, taking into account the user's input in the form of prompts. The procedure is in fact easier to implement than to understand. Only recently, and with some help from advanced physics (Ikeda et al. 2025), has light begun to be shed on its inner workings. Diffusion processes are stochastic, which means that AI results depend heavily on probabilities. It is therefore no surprise that similar inputs can produce different results. This stands in stark contrast to the original concept of photography, in which photographing the same scene repeatedly would yield the same output.

During the diffusion process, data are compressed into an efficient form that also stores the relations between them, creating what is called *latent space*. Many researchers have shown great interest in the topology of latent space and in what it implies for the epistemology of the underlying data. Latent space is only an intermediate step towards creating AI images, and the data at this stage are not meant to be visualized. With special commands, however, this is feasible, producing image instances caught mid-process. This was of special interest to Voto (2022), who studies an artistic video made up of shots of latent space. In it, human faces with mixed characteristics are portrayed as image production gradually moves through iterations of denoising. With varying degrees of success, and with some artistic licence, some of these images could be perceived as "real", that is, as coming from an actual photoshoot.

Voto recognizes that latent space is a step "that usually remains invisible to the human eye, maintaining the promise of a transparent effect of reality" (Voto 2022: 55). Latent space is indeed an intermediate process not meant to be seen – not in order to hide anything, but because compressing and processing data is in no way the same as the final product. "Seeing" data before they are rendered "visible" creates an epistemological void of its own. Yet, with the inner workings of AI still largely a "black box", any conclusion reached may be enlightening. Voto explains that the information in datasets is closer to *capta* than to *data* (Voto 2022: 59), echoing the concerns of others: data are not neutral, either within the dataset or in the processing phase.

In any case, what challenges our senses is our inability to distinguish artificial images easily from photographs in the classic sense. Several events already promote AI-generated artistic content, from beauty pageants (Ronald 2024) to film festivals. The omnipresence of algorithmic content, combined

with the impression it makes on us, may soon normalize the new epistemological status of the artificial. Shaw and Nave (2026) have aptly named this transition *cognitive surrender*: a state of cognition in which AI outputs are adopted with minimal scrutiny. Their research with human participants showed that people favour using AI in decision-making and problem-solving and, moreover, do so with greater confidence despite obvious errors. This mode of cognition contrasts with the traditional modes of cognition that take place inside our brains, whether fast or slow. Shaw and Nave are quick to point out that they do not intend their results as a warning about the dangers of AI, but as "a recognition of System 3's psychological presence" (Shaw & Nave 2026: 50), System 3 being their term for the AI-based mode of cognition associated with cognitive surrender.

Cognitive surrender describes an end state in which the acceptance of the epistemic status of AI images is taken for granted. It may sound similar to *suspension of disbelief*, a literary term for the cognitive distancing required to achieve emotional involvement in a narrative (Chandler & Munday 2020). The latter, however, is voluntary and serves an aesthetic function, whereas the former is involuntary and serves an epistemic one.

To summarize, the probabilistic nature of AI-generated content (images, in our case) makes for a less dependable epistemic affordance. The same probabilistic nature, combined with our readiness to accept it, can reinforce false knowledge and help it spread.[1] Our inability to distinguish the real from the generated is an obstacle to breaking this cycle of reinforcement. Admittedly, however, at least two aspects speak in favour of AI-generated images. First, they do not get everything wrong. The amount of information they get right is remarkable, and as technical limitations are overcome, the situation will probably improve. A prompt asking for a house, for example, will produce an image with doors and windows in the right places rather than placed at random. The training dataset and the corresponding latent space have evidently captured some basic physical laws and social norms correctly, even if only in the sense of how reality usually looks. The results are inductive rather than deductive, but as long as we are aware of this, AI tools can prove epistemologically helpful.

[1] Another detail that shows the probabilistic nature inherent in AI is that early image-generation models produced several outputs at once, usually four, at a time when computational cost was not yet prohibitive. The reason was that, on examining them, users were more likely to find one result suited to their purpose and could easily discard the unsuccessful ones. The significant progress of newer models led to the abandonment of this practice.

Second, AI can be used with equal ease to visualize not what is but what could have been, with images ranging from the imaginary to the surreal and the abstract (for a good example, based on the prompt "fingers kissing", see Zylinska 2024). In this case, the output is closer to art than to engineering. The problem is that, however specific the user's prompts may be, we cannot be certain whether the algorithm is actually taking all the information into account or whether, through some stochastic error, it presents the impossible as possible and vice versa; since we are unable to tell them apart, this could become a source of systematic errors.

The situation becomes even more blurred in areas where reality (that is, physical laws) must coexist with artistic design – which is exactly the case in architecture. Architectural designs must cater to aesthetic preferences as well as practical needs and, above all, they must be feasible. The next section focuses on uses of AI in architecture and presents a set of tools that can bridge this epistemic gap.

5. AI in art and architecture

As explained above, responsiveness and ease of use were among the main reasons for the rapid spread of AI tools. Early adopters experimented for recreational purposes, but corporations soon took over, and the crucial questions of intellectual property and of the nature of human creativity came to the surface. Whether AI is creative or not is beyond the scope of this article. What needs to be made clear is that, for many artistic endeavours, the photorealistic is enough; no special connection to reality is needed. Indeed, art often serves as a vehicle for escaping reality. This is another reason, besides cost-cutting, why AI was quickly adopted in marketing and advertising. Notable examples include campaigns by Coca-Cola (Coca-Cola Company 2025), BMW (Boeriu 2026) and Nutella (theARlab 2026).

Architecture, which combines art with science, cannot afford to limit itself to what merely looks realistic or beautiful: it needs to be buildable. In this respect, AI tools already play a major part in contemporary design. As mentioned above, the Austrian architecture firm Coop Himmelb(l)au has an established AI branch, Deep Himmelb(l)au. To the best of the author's knowledge, this is still an isolated case, but it shows a tendency to take AI seriously in architectural design. A look at its website makes it clear that Deep Himmelb(l)au regards AI as a creative rather than an engineering tool.

This does not mean that there are no AI tools for technical workflows. TestFit can automatically generate site plans to support the building process. Autodesk lists a whole range of AI tools on its website, including "AI-powered generative design tools", "optimization of resources" and "predictive maintenance" (Holmes 2023). Tencent's Voyager is a video generator that can create explorable 3D reconstructions from 2D images. Combined with 3D printers, such tools can fully automate the creation of an object. The collective volume edited by As and Basu (2021) and the extensive monograph by Chaillou (2025) provide recent theoretical and applied examples of the use of AI in architecture. Leach (2025) presents architectural artefacts that were generated with AI from design to construction, such as a chair (Fig. 1a) or the room-sized structure *Cloud Pergola* (Fig. 1b).



(a) (b)

Figure 1: (a) 3D-printed GAN chair (Koh 2020) and (b) *Cloud Pergola* (Andrasek & Juričić 2018), as presented in Leach (2025: 33, 113)

The next section presents some cases in which the gap between AI-generated architectural designs and reality made misuse possible.

6. Beyond the impressive: problematic cases of AI in architecture

This section presents a set of examples in which AI was used with negative consequences because the images lacked epistemological consistency. As is usually the case with any new tool, mistakes and harmful uses are attributed to the users, who are blamed either for a lack of skill or for selfish motives. But since both are to be expected, the absence of mechanisms allowing stricter control is also a cause of such incidents. A combined framework of semiotic, cultural and epistemological criteria is therefore proposed to check the epistemic validity of AI images.

This triadic approach can be broken down into the following components. The semiotic component concerns enunciation: how the image persuades the viewer of its realism, and which details of its framing make the content believable. It addresses the overall visual qualities that make an image believable at first glance, before any closer scrutiny. The cultural component examines the actual content of the image and its composition, and checks them against cultural norms. A well-documented case is the depiction of racially diverse Nazi soldiers by early AI models without any such prompt (Jacobi & Sag 2024). Finally, the epistemological component checks for physical inconsistencies, from relative dimensions to the behaviour of light. The approach is summarized in Table 1.

Table 1: The semiotic–cultural–epistemological framework of validation

Dimension	**Question**	**Example failure**
Semiotic	Is the image believable regarding its realism?	Overly stylized content
Cultural	Does it respect social expectations?	Improbable situations (e.g. water slides in residential zones, lush flora on arid island)
Epistemological	Is it physically plausible?	Impossible lighting, dimensions

The images analysed do not form a sample in the statistical sense; they serve as a demonstration rather than as a test with statistically robust results. The selection criteria were generation by AI, architecture as the central theme and popularity, based on easily accessible metrics (likes, comments, views).

The first image (Fig. 2) comes from the personal Facebook account of a professional architect (since it is a personal rather than a professional profile, no further details can be given). It is a top-view design of a bathroom in photorealistic rendering. A client had given the image to the architect as a guideline, admitting that it had been made with AI, and the architect shared it with some amusement. As it turned out, not only were the measurements completely wrong, but some of the design choices were of dubious practicality.

The problems in Figure 2 are quite obvious. Although there are no misalignments, there are cultural errors (too many basins, cramped spaces)



Figure 2: Design proposed by a client, made with AI (image retrieved from the personal Facebook account of a professional architect)

as well as epistemological ones (distorted perspective on the right, incorrect measurements and relative dimensions). Aesthetically, the image is pleasing and could be mistaken for a real design by anyone willing to overlook these issues.

Still, the question begs to be asked: what made the client think that this image could legitimately help the designer in his work? The possible answers can be narrowed down as follows. First, the client evidently focused on aesthetics rather than technical specifications: the image showed how the bathroom was supposed to "look", not how it was to "be". In other words, the information conveyed was a perceived reality, distorted in relation to the actual one. Second, the client most probably ignored the numerical information or was unable to decipher its meaning. So, apparently, did the AI algo-

rithm, which treated the numbers as decorations that usually adorn designs associated with prompts such as "architectural": the AI is evidently oblivious to their meaning. Third, the result seems to have been so impressive – aesthetically or because of the ease with which it was made – that the client failed to notice its impracticalities (such as the basin inside the shower). This last point appears to be common to other recent applications of AI as well: we seem so enthralled by the responsiveness of AI models that we become forgiving and less critical of their output and its usefulness. A final point is that computational architecture tended to idealize visualization even before the advent of AI, though not at this scale or speed. In short, although the cultural and epistemological flaws were present, only the professional could perceive them; both the client and the AI stopped at the semiotic level of inspection.

This is not an isolated example. Some images become so successful that they go viral on social media (Fig. 3), often leading people to wonder how they could own a similar house or why engineers do not take inspiration from such designs. Although less outrageous than the previous example, such images still make it relatively easy to discern discontinuities. Figure 3 is a refined picture, without distortions, and with fewer clues that reveal cultural slips (the absence of a roof being the most prominent). From an epistemological point of view, the relative dimensions – the chairs are half the height of the door – betray the work of an artificial agent.

Another factor that may explain the popularity of these images is that they express a widespread desire for beautiful housing (or simply for housing). Such images take up space in the collective imaginary, creating expectations that are impossible to meet and even leading to conspiracy theories that beautiful houses are being withheld, when in reality they simply cannot exist. A similar trope concerns small, and thus affordable, houses that seem able to fit rooms and spaces in equally impossible ways; they are often described as having an IKEA aesthetic, because they resemble the brand's use of space in its stores.

The most interesting case, however, is that of exotic houses supposedly located on the Greek island of Santorini, a renowned tourist destination associated with fine living and luxury. The images attracted media attention and led to an article on ellinikahoaxes.gr, a Meta-funded fact-checking site that investigates fake news. The author of the article presented many different examples (Figs. 4, 5 and 6) and pointed out that in at least one case (Fig. 7) the image had been used in the advertising campaign of a travel agency (Vatala 2024), leading potential customers, especially foreign ones, to believe that the place depicted actually exists. If such customers



Figure 3: Example of AI designs circulating through dedicated Facebook accounts. The number of reactions and comments shows how popular such posts become (screenshot taken on 6 September 2024 from the "Tiny House" Facebook account)

booked a package, arrived on the island and discovered that the places depicted did not exist, the resort's reputation could be damaged. Figures 4, 5 and 6 come from the Instagram profiles of individual architects and explicitly mentioned the use of AI, but this information was lost as the images circulated. Figure 7 is obviously AI-generated but of unknown origin (Vatala 2024).

Let us begin with Figure 7, which proved the most controversial. It shows a large water slide with tourists apparently enjoying themselves. Careful inspection, however, reveals the usual technical errors: incorrectly positioned limbs and blurred faces. In other words, no specialized training or method



Figure 4: First AI image of Santorini (reproduced from ellinikahoaxes.gr; original source: Instagram account @LUXREVIVE, retrieved 24 May 2024)



Figure 5: Second AI image of Santorini (reproduced from ellinikahoaxes.gr; original source: Instagram account @vinh.qp, retrieved 24 May 2024)



Figure 6: Third AI image of Santorini (reproduced from ellinikahoaxes.gr; original source: Instagram account @khaled_sadeden_design, retrieved 24 May 2024)



(a) (b)

Figure 7: (a) Fourth AI image of Santorini; (b) the image used for advertising purposes, with personal data blurred (reproduced from ellinikahoaxes.gr; original source unknown)

is needed to spot the epistemological mistakes in the image. This brings us back to the same question: how could it become part of an advertising campaign? This is all the more puzzling given that the use of AI in real-estate fraud is documented to be increasing (Lin 2025). There are two possible, not entirely unrelated, explanations. The first is that these details simply went unnoticed: the image seemed legitimate enough, and no further investigation was deemed necessary. The second is that, even if the errors were spotted, they mattered little to the final outcome. The image remains attractive in itself, and the coloured message embedded in it draws the eye to the information provided – the bargain price – to encourage further action. In both cases, the image is impressive enough to fulfil its purpose.

Examining the other three images shows this reasoning unfolding further. The absence of technical glitches and the masterful editing create a mesmerizing effect: the images are too impressive. This matters because it keeps the viewer in a *sensorial trance*, weakening the will to break the illusion. In other words, the hyperreal fidelity and abundant stylization of the image can inhibit the viewer's critical judgement and facilitate acceptance. This sensorial trance thus marks the boundary between suspension of disbelief and cognitive surrender, discussed above – a point of intersection at which the lowering of scrutiny no longer happens consciously, because of sensory overload. The sensorial trance can be considered a basic component of AI images that pass the semiotic level of the proposed validation framework.

A second, calmer look easily reveals details that culturally aware viewers recognize as wrong. Whether it is a water slide in the middle of a residential area or brightly coloured flora on an arid island in the heat of summer, the actual content of the composition is implausible, however alluring.

Turning to the epistemological component of validation, we quickly find that the more stylistically refined the images are, the more difficult they are to deconstruct scientifically. Figures 4 and 6, however, show harsh lighting with sharp shadows under an otherwise cloudy sky. This is physically impossible, but light breaking through the clouds makes for a more picturesque image and a richer sensory experience.

Another method was proposed by Farid (2017), a specialist in the detection of manipulated images. He suggests using light paths as a clue, since "[a]ll of the light rays, regardless of the scene geometry, should intersect at the same point, whether the light source is nearby or distant" (Farid 2017: 80). Applying the light-path method shows that Figures 4, 5 and 6 are artificial (Figs. 8, 9 and 10). The method is more difficult to apply to Figure 7,

however, because of the lack of clear light paths (in the form of shadows or reflections), which makes a multi-component approach more suitable.



Figure 8: Non-intersecting light paths in Figure 4



Figure 9: Non-intersecting light paths in Figure 5

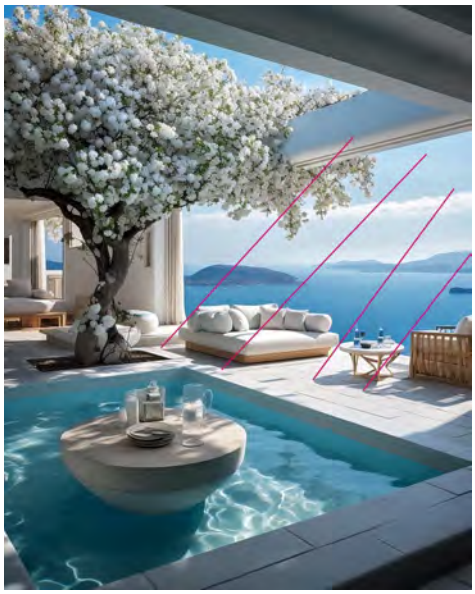

Figure 10: Non-intersecting light paths in Figure 6

7. Conclusions

The rapid spread of generative AI has made the distinction between the real and the merely photorealistic increasingly difficult to draw. Architectural imagery is a particularly interesting field in which to study this shift, because architecture occupies an intermediate position between artistic imagination and scientific feasibility. Unlike purely artistic images, architectural representations are expected to convey information about spaces that could actually be built and experienced. As a result, errors in AI-generated architectural images are not only aesthetic but also epistemological.

The cases examined in this study demonstrate that AI-generated architectural images can be highly convincing even when they contain significant inconsistencies. These inconsistencies were grouped into three categories: semiotic (over-stylization), cultural (situations that conflict with expectations) and epistemological (violations of physical constraints). Examined separately, each category may fail to reveal the artificial nature of an image; together, they provide a more reliable basis for evaluation.

The Santorini example is particularly revealing. The popularity of the images cannot be attributed solely to their technical sophistication. Their success stems from their ability to produce an immersive visual experience that encourages viewers to focus on aesthetic appeal while overlooking inconsistencies. As visual errors become less noticeable in newer AI systems,

such images become harder to challenge and more capable of inducing what this article has called a sensorial trance. In such cases, suspension of disbelief gradually extends beyond artistic appreciation and approaches what recent literature has termed cognitive surrender: the tendency to accept AI outputs with reduced critical scrutiny.

The findings therefore suggest that the central issue is not whether AI-generated images are true or false in an absolute sense, but how appropriate methods for interpreting them can be developed. Architectural AI images should be approached neither exclusively as technical documents nor exclusively as artistic artefacts. They occupy an intermediate epistemic position that requires contextual evaluation. The semiotic–cultural–epistemological framework proposed here offers one possible way of carrying out such evaluations and of reducing the risks associated with misleading or deceptive representations.

More broadly, the study supports the view that the emergence of AI imagery is contributing to a transformation in the way truth, trust and visual evidence are understood. In an environment where photorealism can be generated almost instantly and on a massive scale, credibility can no longer be inferred from appearance alone. The challenge ahead is therefore not merely technological but epistemological: learning how to remain critical observers in a visual culture increasingly shaped by algorithmic production.

References

Alvarado, R. 2023. "AI as an Epistemic Technology". *Science and Engineering Ethics*, Vol. 29, Article 32. Available at: https://doi.org/10.1007/s11948-023-00451-3.

Andrasek, A., B. Juričić. 2018. *Cloud Pergola* [Architectural installation]. 16th International Architecture Exhibition, La Biennale di Venezia. Available at: https://www.labiennale.org/en/architecture/2018 (accessed 20/10/2026).

Anichini, G., B. Kotras. 2024. "Trained Judgements: Artificial Intelligence, Epistemic Tensions and the Production of Scientific Objectivity". *Science, Technology, & Human Values*, Vol *51*(3), 631-663. Available at: https://doi.org/10.1177/01622439241262854.

As, I., P. Basu (eds.). 2021. *The Routledge Companion to Artificial Intelligence in Architecture*. London and New York: Routledge.

Bainss, A. 2023. Exploring the Intersection of AI and Pop Culture. Rephrase.ai. https://dokumen.pub/ai-generated-popular-culture-a-semiotic-perspective-3031547527-9783031547522-x-4816097.html

Bazin, A. 1967. *What Is Cinema?* Vol. 1. (transl. Hugh Gray). Berkeley: University of California Press.

Boeriu, H. 2026. BMW Made an Entire Car Ad with AI. No Camera, No Crew. *BMW BLOG*, 30 April. Available at: https://www.bmwblog.com/2026/04/30/bmw-ix3-generative-ai-campaign/ (accessed 29 October 2026).

Campolo, A., K. Crawford. 2020. "Enchanted Determinism: Power without Responsibility in Artificial Intelligence". *Engaging Science, Technology, and Society*, Vol. 6, 1–19. Available at: https://doi.org/10.17351/ests2020.277.

Caporusso, N., K. Zhang and G. Carlson. 2020. Using Eye-Tracking to Study the Authenticity of Images Produced by Generative Adversarial Networks. In *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*. Istanbul: IEEE, 1-6. Available at: https://doi.org/10.1109/icecce49384.2020.9179472.

Chaillou, S. 2025. *Artificial Intelligence and Architecture: From Research to Practice*. (2nd ed.). Basel: Birkhäuser.

Chandler, D., R. Munday. 2020. *A Dictionary of Media and Communication*. (3rd ed.). Oxford: Oxford University Press.

Chesterman, S. 2024. "Good Models Borrow, Great Models Steal: Intellectual Property Rights and Generative AI". *Policy and Society*, Article puae006. Available at: https://doi.org/10.1093/polsoc/puae006.

Coca-Cola Company. 2025. Coca-Cola | Holidays Are Coming [Video]. YouTube, 3 November. Available at: https://www.youtube.com/watch?v=Yy6fByUmPuE (accessed 20/10/2026).

Durán, J. M., K. R. Jongsma. 2021. "Who Is Afraid of Black Box Algorithms? On the Epistemological and Ethical Basis of Trust in Medical AI". *Journal of Medical Ethics*, Vol. 47, No. 5, 329–335. Available at: https://doi.org/10.1136/medethics-2020-106820.

Event Horizon Telescope Collaboration. 2019. "First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole". *The Astrophysical Journal Letters*, Vol. 875, No. 1, Article L1. Available at: https://doi.org/10.3847/2041-8213/ab0ec7.

Farid, H. 2017. "How to Detect Faked Photos". *American Scientist*, Vol. 105, No. 2, 77–81. Available at: https://doi.org/10.1511/2017.105.2.77.

Ford, M. 2018. *Architects of Intelligence: The Truth about AI from the People Building It*. Birmingham: Packt Publishing.

Helmus, T. C. 2022. *Artificial Intelligence, Deepfakes, and Disinformation: A Primer*. Santa Monica: RAND Corporation. Available at: https://www.rand.org/pubs/perspectives/PEA1043-1.html (accessed 29 October 2026).

Holmes, J. 2023. Vision Setting and Problem Solving: AI in Architecture Is Changing Design. Autodesk, 20 June. Available at: https://www.autodesk.com/design-make/articles/ai-in-architecture (accessed 20/10/2026).

Ikeda, K., T. Uda, D. Okanohara and S. Ito. 2025. "Speed-Accuracy Relations for Diffusion Models: Wisdom from Nonequilibrium Thermodynamics and Optimal Transport". *Physical Review X*, Vol. 15, No. 3. Available at: https://doi.org/10.1103/x5vj-8jq9.

Jacobi, T., M. Sag. 2024. "We Are the AI Problem". *Emory Law Journal Online*, Vol. 74, 1–18.

Joksimovic, S., D. Ifenthaler, R. Marrone, M. De Laat and G. Siemens. 2023. "Opportunities of Artificial Intelligence for Supporting Complex Problem-Solving: Findings from a Scoping Review". *Computers and Education: Artificial Intelligence*, Vol. 4, Article 100138. Available at: https://doi.org/10.1016/j.caeai.2023.100138.

Koh, I. 2020. Discrete 3D-GAN-Chairs… Architecture [Project]. NeurIPS 2020 Workshop on Machine Learning for Creativity and Design. Available at: https://www.aiartonline.com/design-2020/immanuel-koh/ (accessed 29 October 2026).

Lakshmi, D., R. S. Tiwari, R. K. Dhanaraj and S. Kadry (eds.). 2024. *Explainable AI (XAI) for Sustainable Development: Trends and Applications*. Boca Raton: CRC Press. Available at: https://doi.org/10.1201/9781003457176.

Leone, M. 2024. "Semiotics of the Black Box: On the Rhetorics of Algorithmic Images". *Visual Communication*, Vol. 23, No. 3, 426–451. Available at: https://doi.org/10.1177/14703572241247120.

Lin, L. 2025. "GAI-Enabled Real Estate Fraud Schemes: Risks, Prevention and Regulations". *Journal of Technology Law & Policy*, Vol. 29, 97–116.

Matthews, T., I. J. Kidd. 2023. The Ethics and Epistemology of Deepfakes. In Fox, C., J. Saunders (eds.). *The Routledge Handbook of Philosophy and Media Ethics*. London and New York: Routledge, Ch. 13. Available at: https://doi.org/10.4324/9781003134749.

Romele, A. 2022. "Images of Artificial Intelligence: A Blind Spot in AI Ethics". *Philosophy & Technology*, Vol. 35, No. 1, Article 4. Available at: https://doi.org/10.1007/s13347-022-00498-3.

Romero Moreno, F. 2024. "Generative AI and Deepfakes: A Human Rights Approach to Tackling Harmful Content". *International Review of Law, Computers & Technology*, Vol. 38, No. 3, 297–326. Available at: https://doi.org/10.1080/13600869.2024.2324540.

Ronald, I. 2024. The Rise of the AI Beauty Pageant and Its Complicated Quest for the 'Perfect' Woman. *CNN*, 27 June. Available at: https://edition.cnn.com/2024/06/27/style/miss-ai-beauty-pageant-scli (accessed 20/10/2026).

Russo, F., E. Schliesser and J. Wagemans. 2024. "Connecting Ethics and Epistemology of AI". *AI & Society*, Vol. 39, No. 4, 1585–1603. Available at: https://doi.org/10.1007/s00146-022-01617-6.

Scotti, A. 2024. "What to Do with AI Images? Towards an Epistemology of Trust". *I Castelli di Yale Online. Annali di Filosofia*, Vol. 12, No. 2, 43–54.

Shaw, S. D., G. Nave. 2026. Thinking – Fast, Slow, and Artificial: How AI Is Reshaping Human Reasoning and the Rise of Cognitive Surrender. The Wharton School Research Paper. Available at: https://doi.org/10.31234/osf.io/yk25n_v1.

Shin, D. 2026. "Automating Epistemology: How AI Reconfigures Truth, Authority, and Verification". *AI & Society*, Vol. 41, No. 2, 1553–1559. Available at: https://doi.org/10.1007/s00146-025-02560-y.

Smits, J., T. Borghuis. 2022. Generative AI and Intellectual Property Rights. In Custers, B., E. Fosch-Villaronga (eds.). *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice*. Information Technology and Law Series, Vol. 35. The Hague: T.M.C. Asser Press, 323–344. Available at: https://doi.org/10.1007/978-94-6265-523-2_17.

Sontag, S. 2005 [1977]. *On Photography*. New York: RosettaBooks. theARlab. 2026. Nutella AI campaign [LinkedIn post], 2 April. Available at: https://www.linkedin.com/posts/nutella-carousel-activity-7445532907659616256-5vcy (accessed 20/10/2026).

Vatala, E. 2024. Αυτές οι εικόνες της Σαντορίνης είναι προϊόν AI και όχι αληθινές [These Images of Santorini Are AI Products, Not Real]. *Ellinika Hoaxes*, 24 May. Available at: https://www.ellinikahoaxes.gr/2024/05/24/aftes-oi-eikones-tis-santorinis-einai-proion-ai-kai-ochi-alithines/ (accessed 20/10/2026).

Voto, C. 2022. "From Archive to Dataset: Visualizing the Latency of Big Data". *Punctum: International Journal of Semiotics*, Vol. 8, No. 1, 47–62. Available at: https://doi.org/10.18680/hss.2022.0004.

Walton, K. 1984. "Transparent Pictures: On the Nature of Photographic Realism". *Critical Inquiry*, Vol. 11, No. 2, 246–277.

Yampolskiy, R. V. 2020. "Unexplainability and Incomprehensibility of AI". *Journal of Artificial Intelligence and Consciousness*, Vol. 7, No. 2, 277–291. Available at: https://doi.org/10.1142/s2705078520500150.

Zylinska, J. 2024. "Diffused Seeing". *Media Theory*, Vol. 8, No. 1, 229–258. Available at: https://doi.org/10.70064/mt.v8i1.1075.