

Article

Negative Capabilities: Investigating Apophasis in AI Text-to-Image Models

Hannah Lucas 

Newnham College, University of Cambridge, Cambridge CB3 9DF, UK; hal32@cam.ac.uk

Abstract: Through a case study of images generated by Swedish artist Steph Maj Swanson using an AI text-to-image (T2I) model, this article explores the strategy of negative weight prompting in T2I models as a phenomenon of apophasis. Apophasis is a linguistic strategy commonly deployed in texts of mystical theology to express the ineffability of God through negative concepts. In this article, a comparison of apophatic strategies in mystical texts and T2I models is engaged to highlight the mutual benefit of theorising AI with the help of religious theory and concepts. With this, the article builds on previous work on the New Visibility of Religion, enchantment, and post-secularism—especially the research of Beth Singler on religious continuities in representations of AI. Recent work on AI prompt engineering, computational linguistics, and computational geometry is invoked to explain the linguistic processes of T2I models. Poststructuralist semiotics is then employed to theorise the search for the Transcendental Signified in apophatic theology. The article concludes that linguistic theology can help to elucidate technological use cases, subsequently arguing for further dialogue between scholars in artificial intelligence and religious studies, and for a revaluation of religion in the technological sphere.

Keywords: artificial intelligence; religion; apophasis; mysticism; mystical theology; negative theology; semiotics; large language models; Julian of Norwich; medieval literature



Citation: Lucas, Hannah. 2023. Negative Capabilities: Investigating Apophasis in AI Text-to-Image Models. *Religions* 14: 812. <https://doi.org/10.3390/rel14060812>

Academic Editor: Simon Podmore

Received: 28 April 2023

Revised: 16 June 2023

Accepted: 16 June 2023

Published: 20 June 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In September 2022, the Swedish artist Steph Maj Swanson, known online by her handle “@supercomposite”, was experimenting with an AI text-to-image (T2I) generation model when she encountered a strange set of images. Swanson first entered the search term “Brando::–1”—an input called a *negative weight prompt*, which would, in this case, instruct the model to generate an image as maximally dissimilar to “Brando” (i.e., Marlon Brando) as possible. The result was an incongruous logo of black silhouetted buildings on a mostly plain background, with green text overlaid reading “DIGITA PNTICS” (Figure 1). To test whether the model would in turn generate something close to Marlon Brando if “DIGITA PNTICS” was entered as a negative, Swanson then entered “DIGITA PNTICS skyline logo::–1”. Posting her findings on a Twitter thread, Swanson described a human figure in the resulting images: a “devastated-looking older woman with defined triangles of rosacea(?) on her cheeks” (Swanson 2022b). She named this figure “Loab”, on account of the text which accompanied one of the images (Figure 2).

Riffing off the original Loab images, Swanson used the model to recombine Loab with other images, such as a “hyper compressed glass tunnel surrounded by angels [. . .] in the style of Wes Anderson” (Swanson 2022a). The result was a series of increasingly disturbing and macabre scenes, including uncanny doll-like figures dripping blood and even dismembered, screaming children. Amid all of these, the figure of the woman with red cheeks persisted. The experiment invited a number of responses and articles ironically decrying the “demon” Loab, accusing Swanson of summoning an evil spirit from the AI model. As Swanson put it in a tongue-in-cheek tweet: “To clarify for the press [. . .] I have brought a real IRL demon to life. Research has found that demons are real and live inside

of computers. Computers are like little houses for demons and church is like a big house for angels” (Swanson 2022a).



Figure 1. DIGITA PNTICS Logo (Swanson 2022b).

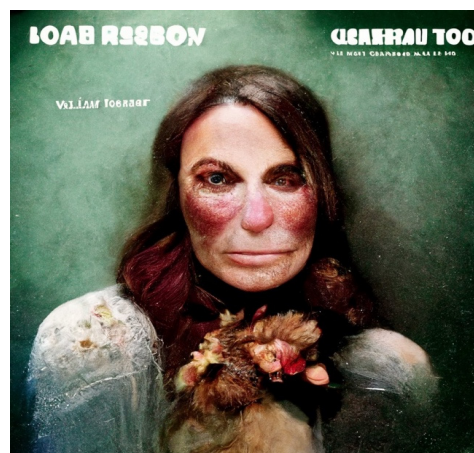


Figure 2. Loab (Swanson 2022b).

Swanson’s invocation of demons and angels here speaks to recent work in the anthropology and sociology of religion and science, focused on the increased visibility of religious imagery and language in AI discourses (Singler 2020a, 2020b; Geraci 2010). The claims (repeated by her online respondents) that Swanson summoned a demon can in part be attributed to the repetitively macabre quality of the images, and in part to the apparently inexplicable emergence of the figure from the negative prompt “DIGITA PNTICS::–1”. Why are the images so consistently gory? Why the appearance of this specific woman? The first phenomenon—the images’ repeated characteristics—is not quite as inexplicable as it might appear. As Paul Tassi (2022) notes in his Forbes article, “Meet Loab, The AI Art-Generated Demon Currently Haunting The Internet”, “The AI is not surfacing Loab images for every negative prompt, what’s being done here is that the *original* image of Loab is being used in combination with new prompts and other art to create these new images. It is true that yes, these are weighted toward being increasingly disturbing, and Loab’s signature features are always retained, which is some interesting AI bias here, but Loab is . . . not an actual demon haunting negative prompts in AI art generators” (emphasis in original). To clarify: it *is* interesting that the characteristics of the Loab character are retained but ultimately not wholly surprising given that the original Loab image was recombined with a series of other images. Regarding the images’ macabre character, Swanson herself points out that the features the AI model identifies with Loab are likely “adjacent to extremely gory and

macabre imagery in the distribution of the AI's world knowledge", which she terms "some kind of emergent statistical accident" (Swanson 2022a).

The following discussion of the Loab case is therefore less interested in the quality of these images themselves, focusing instead on the latter phenomenon: the emergence of the images from Swanson's negative prompting. The article explores this as a process of *apophasis*: the linguistic process of identifying a subject through negating concepts and language. In seeking to elucidate the negative prompt functionality of T2I models, my intention is not to rationalise away the supernatural discourse which attends to this case study. Rather, the article's methodology builds on work in linguistic theology and the philosophy of religion to explore how religious discourses can and do productively intersect with advances in AI. Through a close analysis of the linguistic protocols of T2I models against the backdrop of western mystical theology, the article argues for the value of ongoing dialogue between religious and computational studies.¹ In this, I respond to the work of Beth Singler, who argues that "there is [. . .] no human teleology just as there is no AI teleology predicting a post-human future. What we have [. . .] are symbols and stories", which are "remixed into new shapes that evoke past imaginaries and are made newly visible" (Singler 2020b, p. 15). This article posits that Loab is another instance of such a story, and that linguistic theology can help us better understand this case study and the processes behind this emerging language-based AI technology.

While Singler is occupied with "the stories we tell ourselves about the future of both AI and religion" (Singler 2020b, p. 15), however, this article draws on the "past imaginaries" which are made newly visible in the technological context. By examining the apophatic conventions of both T2I models and works of mystical theology from the Middle Ages, I argue for a backwards glance which not only brings together premodern and modern discourses but demonstrates the continuities between past and present expressions of transcendence. This discussion is intended to at once uplift religious visibility, refute the claims of modernity's disenchantment and secularisation, and call for further dialogue between religious and technological fields. Indeed, the case study will show by example that theories of technology can be greatly illuminated by religious thinking (and vice versa), thus reiterating the view (of, e.g., Bellar et al. 2013; Graham 2013) that the boundary between religion and science, the sacred and the profane, is neither fixed nor impermeable.

First, the article discusses negative weight prompts and their functionality in generative T2I models to clarify the mechanics of these processes and how they can help us understand the Loab phenomenon. The article then lays the groundwork for a comparison of the AI model with linguistic theology by exploring the New Visibility of Religion thesis and discussions around post-secularism. The negative weight prompt is subsequently compared to the strategies of apophasis deployed by a number of mystical texts, exploring the ways linguistic negation can effect cognitive shifts, and how language as a representational tool is theorised in the Christian theological context. Next, I examine the ways these mystical discourses emerge in representations of AI models; comparing the apophatic strategies of mystical theologians and T2I model practitioners, the article here identifies the T2I model as a useful analogue to the theological project, with both exercising the limitations of language. It concludes by arguing for the continued relevance of linguistic theology to emerging questions in AI and machine learning, and why linguists working in this space would do well to consider the ways we have previously theorised the ineffable, on account of their both being bound by the rules and constraints of human implementation and use. In other words, the article contends that theological thinking can help gloss the play of signification which inheres in these generative models.

2. Negative Weight Prompts

Simply put, negative prompts work by providing the T2I model with an instruction as to what it should *not* generate as an image. In a T2I model such as Midjourney (Midjourney n.d.), for example, negative prompts can firstly be added as the "-no" parameter. To generate a surrealist painting without the colour red, one could enter: "surrealist painting

-no red". Alternatively, the negative number method can be used. Here, a negative number (from -0.5 to -1) is added after the element. For instance, to generate a landscape painting with no rivers but more trees, one might enter: "landscape painting::0.5 include trees::1 rivers::-0.5". The higher the weight, positive or negative, the more likely or unlikely it is that the element will appear. With the help of online resources emerging from within the online community, such as Ethan Smith's (2022) *Traveler's Guide to the Latent Space* and Guy Parsons's (2022) *DALL-E Prompt Book*, online practitioners are now widely utilising "prompt engineering" to generate AI art (Oppenlaender 2022; Liu and Chilton 2022; Qiao et al. 2022). Prompt engineering is particularly helpful for developing certain styles of image; for photorealistic images, for instance, negative prompts such as "CG", "fake", "rendered", "artwork", "drawing", or "anime" would instruct the model to stay away from these specific image styles.

In the case of the negative prompt "Brando::-1", this input instructs the model to generate an image based on what it conceives to be maximally dissimilar to the element "Brando". To do so, the model reaches for the furthest point away from "Brando" in its latent space. In layperson's terms, we might gloss latent space as the "knowledge-centre" of the model (though such a term comes dangerously close to assigning agential consciousness to the model itself). Technically, latent space is made up of the relationships between learned data points that a machine learning system has acquired from its training data—a multi-dimensional array of values extracted by the model. The model orders that data such that the most "similar" data points are grouped together, constructing an architecture (a "space") which is defined and shaped by the data that fill it. Because of this architecture's latency, it is difficult to analogise the space or conceptualise it visually.² However, we can think of latent space as multi-dimensional in a couple of senses: firstly, because any two data points can be very close along some axes and far away along others; and secondly, in the sense that latent space can accommodate broad data groupings such as "people", as well as narrower groups such as "children", and single data points which belong to both these categories. These data points, or "embeddings", are gathered together in something akin to a cloud or web. Drilling down into this cloud by using specific prompts can then reveal where these values are located. By identifying mappings between input prompts and embeddings, researchers have been able to alter specific aspects of images—for instance, determining where exactly the data points for "hair" or "mouth" are located for the model to generate images of faces, and operating on them exclusively (Wang et al. 2021). Furthermore, while practitioners still have a relatively limited capability to intervene and edit images directly in this way, advances in generative AI such as Adobe's new Firefly model (Adobe Firefly n.d.) offer new tools which do part of the work for the user. Firefly, for example, combines a model trained on millions of Adobe's own licensed images with an augmented style engine, which pre-encodes instructions to the model to alter the style, colour, tone, lighting, and composition of images.

How the model groups and assigns connections between certain data points depends on how the model is trained. T2I models have been typically trained on large datasets (e.g., Lin et al. 2014) of pre-captioned images (text-image pairs), which the model then organises in its latent space. These datasets are either captioned by human annotators or automatically by AI generators—a process which often exposes models to certain biases. For example, an input prompt to Stable Diffusion (Stable Diffusion n.d.) of "a portrait of a clever scholar who works at a university" will generate images of predominantly white men. Indeed, Luccioni et al. (2023) found that across 96,000 images generated by three popular text-to-image systems (Open AI's DALL-E 2; StabilityAI's Stable Diffusion v 1.4 and v 2), "all three significantly over-represent the portion of their latent space associated with whiteness and masculinity across target attributes; among the systems studied, DALL-E 2 shows the least diversity, followed by Stable Diffusion v2 then v1.4". These findings serve to reiterate that any AI model is only as strong as its training data and its training process.³

The training process cannot always explain how or why certain embeddings or grouped characteristics are related or distinct from others—or why certain images are

generated in response to certain prompts, especially for more obscure prompts such as “Brando:—1”. A closer look at the mathematical makeup of latent space is more helpful. Research in computational geometry has shown that if we are seeking the furthest point from a query point on a three-dimensional shape, it has to be one of relatively few candidates (Pagh et al. 2015). Applying these findings to the Loab scenario, Matthew Skala (2022) explains that whichever query point you start from, the number of possible answers to the question “what is the furthest image from this one?” is much smaller than the set of questions like that that you can ask. Moreover, it is to be expected that relatively few end points will be the same across very different starting points. A helpful geographic analogy to this is a particular set of coordinates in the Pacific Ocean, Point Nemo—also called the “oceanic pole of inaccessibility” because it is the furthest possible place from any landmass. In other words, whichever landmass we start from, Point Nemo is the furthest single point we can get. Just so with the latent space query. Furthermore, Skala suggests that most data are clustered in one area of the latent space, minimising the possible “furthest from here” points available: “If your queries are biased or concentrated in one part of the space, then the number of possible images that ‘furthest from here’ queries could converge on, could be as small as one”. In our seafaring analogy, this is akin to all boats leaving from a few docks on the same coast; they are more likely to arrive at a similar “furthest place”. The upshot is that in the T2I model’s latent space, certain apparently random endpoints may appear (and may appear repeatedly) in response to the model accessing what it conceives to be the furthest possible thing. As Skala points out, this geometrical explanation does not explain the emergence of the precise characteristics of Swanson’s Loab, but it does help us understand how certain images emerge from negative requests.

3. Enchantment and the New Religious Visibility

The following discussion draws a link between this strategy of negative weight prompting and the representational strategies of mystical texts, glossing both using the phenomenon of apophasis. With this, I argue that existing theological frameworks have much to offer theoretical studies of AI models and vice versa. When engaged effectively and rigorously, I propose that this relationship can be mutually illuminating—at once evincing the claim that modernity is not straightforwardly disenchanted, nor that mystical theology is straightforwardly superstitious (or irrelevant in allegedly “secular” contexts). Considered in the context of the language of demon-conjuring invoked by Swanson, the discussion also reiterates Singler’s observation that in public discourses of AI, “religious continuities are often overlaid with overt supernaturalism” (Singler 2020b, p. 13). In highlighting these continuities, I therefore support Singler’s claim (also seen in Bellar et al. 2013) that “the ‘religious’ and the ‘secular’ are fluid categories when it comes to responses to technology and the narratives that emerge around it”. This fluidity at once demonstrates the New Visibility of Religion in modernity as explored by Singler and others, and evinces the need to engage pre-existing imaginaries in order to understand and flesh out current representations.

The demonic language used by Swanson to describe Loab is not unusual in discussions of AI. Singler (2020b) explores such a phenomenon in her analysis of the “AI Creation Meme”, an image of a robotic finger reaching to touch a human finger, which she calls an “AI remix” of Michelangelo’s *The Creation of Adam*. The example provided is David Icke’s version of the AI Creation Meme, which is illustrated with the words “Artificial Intelligence. Summoning the Demon” (Singler 2020b, p. 13). Singler notes the potential inspiration of this version of the meme in Elon Musk’s comment in 2014 at the MIT Aeronautics and Astronautics department’s Centennial Symposium: “With artificial intelligence we’re summoning the demon. You know those stories where there’s the guy with the pentagram, and the holy water, and he’s like—Yeah, he’s sure he can control the demon? Doesn’t work out” (Musk 2014, quoted in Singler 2020b, p. 13). However, Singler also points out the complicating factor of Icke’s “conspirituality”—his “well-documented belief in the ontological reality of non-human and malevolent, transdimensional and interstellar beings

that he calls ‘demons’—which may render his comment less metaphorical and “more theological” than Musk’s. She concludes that both Icke and Musk are employing religious imagery in public discourses around AI and subsequently raising the *visibility of religion* in these spaces (p. 14).

In the wider community of practitioners working with T2I models, too, the invocation of supernatural language is evident. As Oppenlaender’s taxonomy of prompt modifiers for T2I models exemplifies, one type of modifier commonly used in the AI art community is that of the “magic term”: “terms which introduce randomness to the image that can lead to surprising results” (Oppenlaender 2022, p. 8). Oppenlaender gives the example of Twitter user John David Pressman (@jd_pressman), who added the magic term “control the soul” to the prompt “orchestra conductor leading a chorus of sound wave audio waveforms swirling around him on the orchestral stage” (Pressman 2021). The term was added to—in his words—produce “more magic, more wizard-ish imagery” (Pressman 2021, quoted in Oppenlaender 2022, p. 8). Magic terms, Oppenlaender notes, can refer to “subjects that are only distantly related to the main subject of the prompt, or they can refer to non-visual qualities, such as the sense of touch (somatosensory), sense of hearing (auditory), sense of smell (olfactory), and sense of taste (gustatory) (e.g., ‘feed the soul’ and ‘feel the sound’)” (Oppenlaender 2022, p. 8). In other words, these are prompts which gesture beyond the realm of the visual and into the realm of the affective in order to open up more generative possibilities.

The signifier “magic” in this context draws from supernatural and religious-adjacent imaginaries, thus also raising the visibility of religion as in Singler’s example. Indeed, when paired with the invocation of the “soul” and attendant supra-visual characteristics, the above description shares elements with the language used in mystical texts.⁴ This is not, moreover, a gesture towards irrationality, since the functional power of the prompt is ontologically real; in other words, the prompt *works*. In the prompts themselves, supernatural language is used to enhance or elevate the model’s output into a more transcendent sphere—thus drawing on pre-existing embeddings within the model which encode notions of “soul” in its latent space. Furthermore, in the description of this function as “magic”, the same pre-existing imaginaries of symbols and stories are being employed—imaginaries which understand transcendence as both accessible and even rational. Early medieval conceptions of magic, for example, should not be understood as irrational or nonrational, since as Richard Kieckhefer has argued, “to conceive of magic as rational was to believe, first of all, that it could actually work [. . .] and, secondly, that its workings were governed by principles (of theology or of physics) that could be coherently articulated” (Kieckhefer 1994, p. 814). Whether this magic was divinely or demonically instigated was a matter of great theological and socio-political concern in the premodern period, since both could manifest in comparable ways. As the above examples illustrate, such distinctions are blurred even further in the AI hermeneutic, where religious continuities and supernaturalism are not only overlaid, but interwoven into a new fabric of technological transcendence.

Such complex invocations of supernatural, religious, and quasi-religious ideas in AI discourses serve to reiterate Jason Josephson-Storm’s statement that “it seems hard to assert that we live in a straightforwardly disenchanted world” (Josephson-Storm 2017, p. 34). *Disenchantment* is a term borrowed from Friedrich Schiller by Max Weber and taken up by scholars such as Ernest Gellner (1979) to describe the secularisation of modern society. This secularisation thesis has been since challenged by scholars objecting to both this notion of disenchantment, and that we have ever been “modern” at all (Latour 1993). In *The Myth of Disenchantment: Magic, Modernity, and the Birth of the Human Sciences* (2017), Josephson-Storm takes the former position, exploring the ways that enchantment has persisted and continues to persist across the domains of the sacred and the profane. As Singler has subsequently explored, these representational strategies serve to highlight “continuities of mythologizing and enchanted ways of conceptualizing the relationship between the human and the non-human other” (Singler 2020b, p. 2).

Singler's term "continuities" stresses the diachronicity at play in such discourses, for which the present study is also arguing. This diachronic perspective understands religious imaginaries as pervasive throughout history, and now emerging in new and enlivening ways. This "New Visibility of Religion" thesis thus challenges the narrative of *re-enchantment* following rational, post-Enlightenment discourses—or of the *post-secular*, an expression similarly rooted in the notion that modernity is characterised by secularism or disenchantment. As Michael Hoelzl and Graham Ward put it in the introduction to their volume, *The New Visibility of Religion: Studies in Religion and Cultural Hermeneutics*, "we are [not] experiencing a re-emergence of religion, but a new visibility—a visibility that is far more complex and nuanced than the simple re-emergence of something that has been in decline in the past but is now manifesting itself once more" (Hoelzl and Ward 2008, p. 1). In other words, they propose a "*new awareness* of religion" and new recognitions of "forms of religious beliefs and their manifestations that we have never encountered before" (p. 2). Instead of arguing that we are somehow *beyond* or *post-* an era of disenchantment, this position therefore makes space for diachronic understandings of spirituality and religion.

In the discussion below, I show how the language of enchantment deployed in the above examples is but a surface-level expression of the possible continuities which can be drawn between religious and technological imaginaries. The following exploration of apophasis in texts of mystical theology highlights the ways this linguistic strategy opens up meaning beyond a literal sign–signification relationship—a strategy analogous to that of the negative weight prompt. In drawing this analogue, the discussion therefore extends the conversation about the New Visibility of Religion and post-secularism, suggesting that religious symbols and stories are not only applicable in *describing* or narrativising technological phenomena, but are profoundly helpful in *theorising* the linguistic processes of these use cases. The intended corollary is a revaluation of religious texts and theory in the technological sphere.

4. Mysticism and the Transcendental Signified

Given the origins of the word *apophasis*—from the Greek *apophatikos*, "denial" or "negation"; literally *apophanai* "to speak off" or "to speak away" (*apo*) from—immediate parallels of the rhetorical strategy with the negative prompt process are clear. However, further elucidation is needed to theorise a connection between theological and AI applications. In what follows, I explore two approaches to apophatic language in texts of mystical theology to better understand linguistic expression, its constraints, and how these constraints play out in T2I models.

In the spiritual context, the apophatic theologian wrestles with a fallen language which necessarily anchors itself in the world of signifiers. To attempt to speak about God is an attempt to break free from this prison-house of language, entering into what Vincent Gillespie and Maggie Ross have called "the flickering play of signification" (Gillespie and Ross 2011, p. 277), where absence—God's ineffability—extends the linguistic project indefinitely. In other words, in the face of God's ultimate inaccessibility, the theologian is condemned to speak *around* or *away* from what God is, either via explicit expressions of what God is *not*, or by apophatic strategies of cognition which seek to collapse language entirely. Like metaphor, apophasis breaks down the dyadic sign–signification relationship of a literal expression, instead opening up words to any number of abstract meanings or signifieds. In other words, where in a literal expression $x = x$, in a metaphor $x = y$, where the qualities of y are transferred onto x . Apophasis meanwhile functions as $x = -y$, where the qualities of y are not capable of describing x , which is God. Both metaphor and apophasis are powerful in their capacity for conferring multiple, if not infinite, meanings. Precisely which meanings are received depends partly upon the context of the words, and in part upon the subjectivity of the reader or interpreter. In opening up sense in this way, apophatic language seeks to redirect consciousness from the world of signs into a deeper, non-conscious mode of awareness.

The ultimate goal of this linguistic play can be described in post-structuralist terms as the search for the “Transcendental Signified”, which Jacques Derrida writes of that which “in and of itself, in its essence, would refer to no signifier, would exceed the chain of signs, and would no longer itself function as a signifier” (Derrida 2000, p. 242). The Transcendental Signified is, in this proposition, “a concept independent of language” (Derrida 1981, p. 20)—a universal or absolute truth, a ground of meaning to which all signifiers ultimately refer. God is one such Transcendental Signified, as he cannot be represented by any signifier, yet as the ground and creator of the world of signs all signifiers ultimately refer back to him. In line with the logocentrism of the poststructuralists, Derrida famously rejects the myth of the Transcendental Signified, instead arguing that meaning is always deferred onwards in an infinite chain of signification—a play of *différance* which refuses any fixed or ultimate meaning. Therein lies the divergence between theological semiotics and poststructuralism, with the former claiming an ineffable ground of meaning to all signification, and the latter claiming relativism within language (that there is no such ground, and that context is all there is).

Nevertheless, poststructuralist semiotics has long been a critical conversation partner with linguistic theology, since the theological framing of human language as ultimately limited in signifying God does not negate the poststructuralist claim to language’s arbitrariness. Put another way: both theology and poststructuralism see language as limited. As Jeffrey Ringer notes, “The mystics [. . .] held a strong belief in the limitations of language, a belief that aligns them with several aspects of poststructuralist thought” (Ringer 2003, p. 4). Ringer gives the example of Walter Hilton’s *Scale of Perfection* (c. 1380–1396) as a treatise on the contemplative search for the divine “in a way that parallels what Jacques Derrida in *Of Grammatology* and Jacques Lacan in ‘The Instance of the Letter in the Unconscious’ centuries later would say about language—namely, that the signified is endlessly deferred” (p. 7). He notes that Augustine, too, propounds a definition of God in *On Christian Doctrine* which corresponds to notions of “lack” in the symbolic order of Lacan. Theologically speaking, the fallenness of human existence renders worldly signification always and essentially imperfect, and thus incapable of ever reaching God. This renders the task of the contemplative a never-ending project of seeking, possibly or possibly not attaining some understanding, and seeking some more—with true fulfilment only available upon the bliss of the beatific heavenly vision.

Contemporary contemplative writings to Hilton articulate this theological semiotic with varying emphases on language’s potential for successfully signifying its transcendental ground. For the anonymous author of the *Cloud of Unknowing* (c. 1375), for example, language (and thus, cognition: *intellectus*) must be ultimately negated for the contemplative to reach the divine. The goal of contemplation as they have it is to “forget everything”, to place a cloud of forgetting “between you and everything in creation” (Anonymous 2009, p. 19). Only by doing this, the author argues, will the contemplative come to terms with the “cloud of unknowing” which exists between humankind and God. The treatise proceeds to offer various cognitive strategies to the reader which will help them toward this goal, one of which is focusing the scattered mind on one small word: “Select a little word of one syllable, not two” (p. 24), writes the author. “*God* or *love* works well. Pick one of these or any other word you like, as long as it is one syllable. Fasten it to your heart. Fix your mind on it permanently, so nothing can dislodge it”. With this single-pointed meditation, the contemplative is to refuse intellectual analysis: “tell [your thoughts] that you refuse to analyze the word, that you want your word whole, not broken into pieces”, the *Cloud*-author writes. The paradox is that by focusing on and repeating the word, the word’s everyday associations of meaning will begin to dissipate, and the signifier will split asunder, opening up space for new meaning and new interpretation. Eventually, in the *Cloud*’s model, even language will fall away in favour of “a blind gazing at the naked being of God alone” (p. 28).

Yet the paradox of the Christian salvation story means that despite redemption—and access to the Transcendental Signified—being always deferred into the future, the death of

Christ has *already* redeemed humanity on earth. Mystical theologian Julian of Norwich provides an account of this eschatology in her visionary text, *A Revelation of Love* (c. 1373–1416). Here, she describes the human soul as split into “substance” and “sensuality”, the former of which is always united to God and in which he dwells. The latter is the experience of human bodiliness, which will lead humanity into sin—and which Christ has redeemed by taking on human form. The contemplative’s lifelong practice of devotional seeking is intended to draw these two parts back together in unity, restoring that which has slipped into unlikeness back to the likeness of God.⁵ This salvation story works on account of multiple concurrent temporalities, where human time coexists with the nonlinear time of the divine. As such, humanity is already redeemed by the historical Christ and always in the process of being redeemed by human cooperation. The consequence for a Christian semiotics is that the world of signs is *both* limited in its capacity to signify God *and* capable of some expression of divinity on earth.

In stressing the incarnational aspect of the Christian story—an affective focus on the figure of Christ and his suffering—Julian of Norwich’s theology is therefore optimistic as well as apophatic, positive as well as negative. As Grace Jantzen points out, where the *Cloud*-author after pseudo-Dionysius would write that “no man can think of God himself”, Julian does believe that “we can have though never exhaustive knowledge of God, at least that kindergarten ABC variety of real communion with him which is the preliminary to the fullness of knowledge which she believes awaits us when we see him face to face” (Jantzen 2000, p. 94). As Julian puts it, “Also in this marvelous example I have techyng with me, as it were the begynnyng of an ABC, wherby I may have sum vnderstandyng of our lo[r]dis menyng” [in this marvellous example I have teaching with me, as it were the beginning of an ABC, whereby I might have some understanding of the lord’s meaning] (Julian of Norwich 1993, p. 79). In other words, Julian understands incarnated signs—including language, and even images—as helpful hermeneutic tools for the contemplative project. Indeed, her own revelation begins by her focusing her attention on a crucifix before her, a strategy of affective piety which leads with compassionate feeling and imagination of Christ’s suffering. For Julian, then, our essential connection to God, paired with Christ’s incarnation and redemptive death, inheres within human life on earth the potential for some understanding of God. Even while we are constrained by the prison-house of our fallen language, Julian suggests that during our tenancy, we might make this house as homely as possible.

In seeking to access the Transcendental Signified, the mystical theologian must therefore chase the play of meaning within the world of signs. For the *Cloud*-author, it is by dwelling in the space of *not-knowing* that the ineffable divine may be witnessed, by stripping out language entirely. For Julian, meanwhile, her incarnational emphasis results in a linguistic theology wherein language is meaningful in its *inability* to express the divine as much as its ability. Cristina Cervone has called this “supereffability”, where “the sought thing cannot be comprehended fully at once in its own nature but must be understood by means of something-it-isn’t, expressed in language-it-isn’t” (Cervone 2012, p. 5). Images and objects are invoked by Julian only to slip away, like the blood which drips like rain from the eaves of a house, spreading like the scales on a herring, and all of creation collapsing into a little thing in the palm of her hand. Her language shimmers, flickers, and plays with signification, continually expanding the web of meaning without ever landing upon one specific sign or signified. This linguistic strategy reflects, moreover, the perpetuity of the contemplative process itself: as Julian writes of her own showings, “and thus I saw him an sowte him, and I had him and I wantid hym” [and thus I saw him and sought him, and I had him and I wanted him] (Julian of Norwich 1993, p. 15). Language gestures ever onward, but crucially, Julian finds meaning *in* this continuous seeking: “sekyng is as good as beholdyng” (p. 16), she insists, a statement which reiterates the necessity of worldly signifiers in seeking the Transcendental Signified which is God. Put simply, meaning is found in the journey, even while the destination remains ultimately and eternally out of reach.

5. Closed and Non-Closing Apophasis

Given their essentially linguistic character, T2I models can be understood as operating within equivalent constraints as the semiotic frameworks described above. That is, the models exemplify in microcosm the same prison-house of language with which all apophatic or mystical writers must contend. This section of the paper explores such a connection to mutually illuminate both the ways that apophatic thought and writing can clarify the protocols of T2I models, and how these models might in turn shed light on the organisation and generation of meaning in linguistic processes such as apophasis. It concludes that neural net models offer fertile ground for further investigations into the connections between language models and theology, and in turn for wider conversations between religious studies and computational linguistics.

The architecture of generative T2I models' latent space offers an illuminating paradigm for discussions of transcendence, modelling digitally the so-called play of signification discussed by the aforementioned poststructuralist and theological thinkers. The model's latent space is an infinite ground of dimensions into which the recorded data can extend. We might therefore think of latent space as the model's discursive context—the ground of meaning behind linguistic expressions. The architecture of this space remains contingent upon the data within it, and so it does not represent an unshifting ultimate truth as conceived in the theological context; the signifiers recorded as data will always determine the shape of this ground. Nevertheless, this space is significant in that it functions as an otherwise inaccessible background to expressions of meaning. As already outlined, prompt queries instruct the AI model to generate images from this space, where an inputted query will access the data associated with this expression. We might understand this as the sign–signification process in action, where input (prompt) and output (image) are mediated by the model's understanding of the prompt's signifying properties. This is, however, a less deterministic process than it might seem at first glance, since each prompt query can lead to slightly different images—as demonstrated by Midjourney's four-tile interface, and the fact that the same prompt will produce slightly different images when inputted multiple times. Against a diffuse background of noise, T2I models will always attempt to pin down a small number of meanings, or one meaning, from the inputted query—but will not always succeed.

There are, moreover, discernible differences between the ways some models process language and human linguistic processing—differences which further complicate the notion of literal sign–signification mapping. As Royi Rassin et al. have recently shown, OpenAI's DALLÉ-2 model “sometimes depicts both senses of nouns with multiple senses at once; and [. . .] a given word can modify the properties of two distinct entities in the image, or can be depicted as one object and also modify the properties of another object, creating a semantic leakage of properties between entities” (Rassin et al. 2022, p. 1). In other words, the model struggles with homonyms, realising multiple meanings at once in one image. The team notes that in human language processing, once the interpretation of a specific word is realised it will not be re-used in the same situation with a different meaning. In DALLÉ-2, however, they found that words and phrases are often interpreted as both meanings at once, with certain words which are otherwise grammatically separate “leaking” into the interpretation to realise multiple senses. One example they provide is the word “bat”, which, when the model is given the prompt “a bat is flying over a baseball stadium” (p. 1), is realised as two entities: as both an animal and a piece of sports equipment. This extends even further into what Rassin et al. call second order concept leakage, where the prompt “A tall, long-legged, long-necked bird and a construction site” will still generate an image including both a crane (bird) and a crane (machine) (p. 4). The study concludes that “DALLÉ-2 does not follow basic principles of symbol-to-entity mapping in language” (p. 5), demonstrating the complexity of the sign–signification process in T2I models, and the sheer multiplicity of discursive meaning which exists in the model's latent space.

Already, then, we know that models can behave in ways which interrupt literal sign–signification mapping. As Rassin et al. (2022) highlight, generating one specific signified

from the inputted signifier is not always possible, giving rise to surprising and unintended outputs. Yet strategies such as negative weight prompts can further subvert the positivism of these models. Just as “magic prompts” introduce seemingly random image attributes to images (Oppenlaender 2022, p. 8), negative prompts generate something beyond that which is encoded as literal signification. By instructing the model to generate the thing *most dissimilar* to a specific element or embedding, the negative prompt introduces elements otherwise inaccessible, difficult to access, or at the very least, *unlikely to be accessed*, by other means. I return here to the notion of the furthest possible point query; conceptualised spatially, we can understand how certain elements which are very unlike many other elements will act as outliers against the clustered groupings of similar elements. Because of the lack of likeness in these elements, they will be difficult to access directly via a positive prompt query. By reaching for the furthest possible thing, a negative prompt circumvents this distance through the invocation of opposites.

In this sense, the negative prompt process does perform a kind of apophasis, deferring meaning to the model’s perceived negation of the thing. Yet to again invoke Frederic Jameson’s term, this process remains confined by the prison-house of language; reaching into its latent space, the model can generate an image only using data points it has already encoded from image captions. Put simply, the model cannot generate signifieds behind or beyond its encoded data. To attempt to reach beyond this data—to break open the symbol and reach into the space of silence—is impossible, just as it is impossible for the mystical writer to represent silence in speech. Nevertheless, this does not negate the comparison between mystical apophasis and the negative prompting process. For even in the theological context, where the mystic believes the Transcendental Signified to exist, just as they believe that they will be able to access divine knowledge in heaven, they must accept the limitations of worldly language. As such, the nature of the mystical text is that it can only ever lead the reader to the edge of signification—to the lip of the apophatic realm. Hence the shimmer and flicker of symbols which proliferate in Julian of Norwich’s *Revelation*, for example, where multiple meanings are invoked at once, and slip away just as they begin to take shape; meaning is deferred endlessly in these attempts to represent the ineffable, leading the reader into the outermost limits of the “latent space” of their subconscious. Similarly, the process of negative prompts can only instruct the model to access the edges of unlikeness in the space of its learned data.

The difference between these applications of apophasis resides in their respective teleologies—the final destinations where each must necessarily land. Mystical apophasis asks the reader to project themselves into the world of signification, to lead them into the depths of the web of discourse, to one image and then another and then another, a “game of mystical hide and seek”, as Gillespie and Ross (2011, p. 277) put it. This game is never-ending, with the contemplative seeker directed ever onwards. T2I models, meanwhile, necessarily close down potential meanings to access the most typical data point in response to the prompt query; they *must* produce an image or set of images as an output. Moreover, as the furthest neighbour theory makes clear, there is a limited number of furthest point queries that the model will converge upon. We can, then, distinguish between these two applications as *closed* and *non-closing* apophatic strategies: where the negative prompt instructs the AI model to generate an image from the “least similar” data to the input, the mystical text seeks to open up meanings to “everything but” the symbol itself. Where one reaches for a furthest data point, the other expands language to breaking point by dissolving one meaning into a sea of signification.

However, despite the closed apophasis of the AI model, wherein meaning can only ever be deferred so far in the direction of an essentially positivist output, this does not negate a reading of this process as potentially transcendent. For an apophatic theologian like the *Cloud*-author, who advocates the total negation of intellectual knowledge in pursuit of the divine, the recourse to images which occurs in T2I models would certainly be a failure of apophasis. Yet for Julian of Norwich, who insists that there is a possibility of accessing God through worldly language, any representational tool is considered to be helpful in

this pursuit. From this perspective, all we have is language, for we are human and exist in the world of signs. The task of the apophatic theologian is to work within this totality, to seek into new and surprising meanings from whichever direction possible—including the sideways glance of negative language. What’s more, an image can be as helpful a focal point as a syllable in the one-pointed strategies of meditation advocated by contemplative texts—as *A Revelation of Love* evidences, when Julian is encouraged to focus upon the crucifix before her by the priest who attends to her: “I have brought to you the image of thy maker and saviour. Look thereupon and comfort yourself therewith”.⁶

With this in mind, a Julianist approach to apophasis would conceive of these AI models as simply another tool in humankind’s toolkit for seeking into something transcendent. For just as mystical language draws attention to the limitations of language, the closed apophasis of T2I models leads us up against the edges of the discursive web, demonstrating the constraints of linguistic strategies of difference. Whether or not we can reach beyond the symbol itself, it is surely this view from the edge that changes the perspective of the looker; meaning is to be found in the *incapacity* of language as much as its capacity. For what does the strange and uncanny emergence of an unexpected image such as Loab signify other than the strange uncanniness of language itself? The negative prompt functions, in this sense, like a spotlight on the apparent arbitrariness of signs. In generating an image from the most dissimilar possible data points, the model casts the symbol out of literal sign–signification logic, reminding us that a dimension of meaning exists beyond that which is positively accessible. Loab is but one example of this: that if we continue to defer meaning through negative prompts, we will enter a space of signification not ordinarily inhabited by everyday language.

6. Conclusions

The process of negative weight prompting in generative T2I models has been explored in this article as a phenomenon of apophasis, which defers meaning beyond literal sign–signification relations and into the periphery of the model’s system of encoded signs. The article compared this process to the linguistic apophasis deployed by a number of mystical texts. Differences emerged in these applications of apophasis, culminating in the respective teleologies of the mystical text and the generative T2I model, where one directs the reader’s cognition onward, and the other is confined to a positivist output. I distinguished these as *non-closing* and *closed* apophasis, respectively. Nevertheless, as Julian of Norwich’s thesis of the “ABC” understanding of the divine through language shows, such confines should not necessarily be conceived as limiting access to the transcendent—for even images can be used as points of departure for further deferrals of meaning when recombined with the practitioner’s imagination and implementation. This is also demonstrated by Swanson’s recombination of the original Loab image with further images, producing an entire series of images around the Loab query.

By comparing the linguistic protocols of apophatic theology and generative AI models, the article also contributes to the New Visibility of Religion, arguing for the mutual benefit of theorising AI with the help of religious frameworks. The article is not intended to reinforce religious representations of AI models that would claim a God in the machine, or a demon, or even that negative prompt weights are “mystical” in the sense of offering access to one or many divinities. Indeed, the focus on apophasis discusses the transcendent only indirectly; any claims that the linguistic strategies can or cannot ontologically express something divine or supernatural ultimately depend upon the faith of the practitioner. By limiting the discussion to the linguistic expression itself—to the incarnated system of signification or world of signs—the article instead concluded that comparable representational limits occur in mystical texts as emerge in AI models. Just as the mystic grapples with the web of signification in attempting to represent the ineffable, turning language upon itself to test its own limits, the T2I model necessarily operates within the totality of signification that governs the discursive realm.

It remains to reiterate the conclusion that T2I models provide fruitful analogues for investigating linguistic strategies of negation, with recent advances in AI offering new and productive ground for linguistic philosophers and theologians alike interested in the limits of language. Indeed, I would go so far as to state that a dialogue between philosophical–theological studies and computational science is an urgent next step in the theorising of language, especially that which seems to draw near to the ineffable. Understanding how the ineffable has been theorised through history is one starting point for this conversation and highlights the continued applicability of these frameworks into the present day, thereby providing new insights into the play of signification in language mystical, medieval, and modern.

Funding: This research received no external funding.

Acknowledgments: With many thanks to Paul Röttger (Bocconi University) for his comments on drafts of this paper and his insights into T2I models. Thanks also to the article’s anonymous readers for their feedback. A version of this paper was presented at *The International Society for Science and Religion’s* conference, ‘Artificial and Spiritual Intelligence’ (20–23 June 2023).

Conflicts of Interest: The author declares no conflict of interest.

Notes

- ¹ My focus throughout the article is on Anglophone Christian theology, but it should be noted that one challenge for further work in this area will be the lack of models trained on languages other than English. For a project which engages global mystical literature with a generative text model, see [Mohamed \(n.d.\)](#).
- ² For a visual representation of the latent space of number representations/embeddings in the MNIST dataset as generated in TensorFlow, see [Anderson \(2022\)](#).
- ³ Google’s Imagen model is presented as an example of where text embeddings from large language models pretrained on text-only corpora are remarkably effective for text-to-image synthesis when compared to other methods ([Saharia et al. 2022](#)).
- ⁴ The anonymous author of *The Cloud of Unknowing* (c. 1375), for example, instructs the reader to “Cultivate self-control of body and soul”, and describes the will as “that strong architect of your soul” ([Anonymous 2009](#), pp. 13, 106).
- ⁵ Julian writes that “thow the soule be ever lyke to God in kynde and substance, restorid be grace, it is often onlyke in condition be synne on manys partye” [though the soul be ever like God in nature and substance, restored by grace, it is often unlike in condition by the sin on man’s part] ([Julian of Norwich 1993](#), p. 60).
- ⁶ For more on affective devotion in this period, see [McNamer \(2010\)](#).

References

- Adobe Firefly. n.d. *Adobe.com*. Available online: <https://www.adobe.com/sensei/generative-ai/firefly.html> (accessed on 20 April 2023).
- Anderson, Martin. 2022. What is the Latent Space of an Image Synthesis System. *Metaphysic.ai*, November 17. Available online: <https://metaphysic.ai/what-is-the-latent-space-of-an-image-synthesis-system/> (accessed on 20 April 2023).
- Anonymous. 2009. *The Cloud of Unknowing with the Book of Privy Counsel: A New Translation*. Translated by Carmen Acevedo Butcher. Boston and London: Shambhala.
- Bellar, Wendi, Campbell Heidi Cho, Kyong James, Terry Andrea, Tsuria Ruth, Yadlin-Segal Aya, and Ziemer Jordan. 2013. Religion in Internet Memes. *Journal of Religion, Media and Digital Culture* 2: 1–39.
- Cervone, Cristina Maria. 2012. *Poetics of the Incarnation: Middle English Writing and the Leap of Love*. Philadelphia: University of Pennsylvania Press.
- Derrida, Jacques. 1981. *Positions*. Translated by Alan Bass. Chicago: University of Chicago Press.
- Derrida, Jacques. 2000. Semiology and Grammatology: An Interview with Julia Kristeva [1968]. In *The Routledge Language and Cultural Theory Reader*. Edited by Lucy Burle, Tony Crowley and Alan Girvin. London and New York: Routledge, pp. 241–48.
- Gellner, Ernest. 1979. *Spectacles and Predicaments: Essays in Social Theory*. New York: Cambridge University Press.
- Geraci, Robert. 2010. *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*. Oxford: Oxford University Press.
- Gillespie, Vincent, and Maggie Ross. 2011. The Apophatic Image: The Poetics of Effacement in Julian of Norwich. In *Looking into Holy Books: Essays on Late Medieval Religious Writing in England*. Turnhout: Brepols, pp. 277–306.
- Graham, Elaine. 2013. Manifestations of the Post-Secular Emerging Within Discourses of Posthumanism. Paper Presented at the “Imagining the Posthuman” Conference at Karlsruhe Institute of Technology, Karlsruhe, Germany, July 7–8. Available online: <http://hdl.handle.net/10034/297162> (accessed on 20 April 2023).

- Hoelzl, Michael, and Graham Ward. 2008. *The New Visibility of Religion: Studies in Religion and Cultural Hermeneutics*. London and New York: Continuum.
- Jantzen, Grace M. 2000. *Julian of Norwich: Mystic and Theologian*. London: SPCK.
- Josephson-Storm, Jason. 2017. *The Myth of Disenchantment: Magic, Modernity, and the Birth of the Human Sciences*. Chicago: University of Chicago Press.
- Julian of Norwich. 1993. *A Revelation of Love*. rev. ed. Edited by Marion Glasscoe. Exeter: University of Exeter Press.
- Kieckhefer, Richard. 1994. The Specific Rationality of Magic. *The American Historical Review* 99: 813–36. [PubMed]
- Latour, Bruno. 1993. *We Have Never Been Modern*. Cambridge: Harvard University Press.
- Lin, Tsung-Yi, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. 2014. Microsoft COCO: Common Objects in Context. In *Computer Vision—ECCV 2014*. Lecture Notes in Computer Science, vol. 8693. Cham: Springer, pp. 740–55.
- Liu, Vivian, and Lydia B. Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. Paper presented at the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22), New Orleans, LA, USA, April 29–May 5; New York: Association for Computing Machinery, pp. 1–23. [CrossRef]
- Luccioni, Alexandra Sasha, Christopher Akiki, Margaret Mitchell, and Yaacine Jernite. 2023. Stable Bias: Analyzing Societal Representations in Diffusion Models. *arXiv arXiv:2303.11408*. [CrossRef]
- McNamer, Sarah. 2010. *Affective Meditation and the Invention of Medieval Compassion*. Philadelphia: University of Pennsylvania Press.
- Midjourney. n.d. *Midjourney.com*. Available online: <https://www.midjourney.com> (accessed on 20 April 2023).
- Mohamed, Kiara. n.d. Digital Mysticism (DM). *Axisweb.org*. Available online: <https://www.axisweb.org/article/digital-mysticism-dm/> (accessed on 20 April 2023).
- Musk, Elon. 2014. Tesla's Elon Musk: We're "Summoning the Demon" with Artificial Intelligence. *YouTube.com*. Available online: https://www.youtube.com/watch?v=Tzb_CSRO-0g (accessed on 20 April 2023).
- Oppenlaender, Jonas. 2022. A Taxonomy of Prompt Modifiers for Text-To-Image Generation. *arXiv arXiv:2204.13988*. [CrossRef]
- Pagh, Rasmus, Francesco Silvestri, Johan Sivertsen, and Matthew Skala. 2015. Approximate Furthest Neighbor in High Dimensions. In *Similarity Search and Applications SISAP 2015*. Lecture Notes in Computer Science vol. 9371. Edited by Giuseppe Amato, Richard Connor, Fabrizio Falchi and Claudio Gennaro. Cham: Springer, pp. 3–14.
- Parsons, Guy. 2022. *The DALL-E 2 Prompt Book*. Available online: <https://dallery.gallery/wp-content/uploads/2022/07/The-DALL%E2%82%B7E-2-prompt-book-v1.01.pdf> (accessed on 20 April 2023).
- Pressman, John David. 2021. Twitter thread: King of the Podcasts Summons His Subjects [. . .]. *Twitter.com*, November 7. Available online: https://twitter.com/jd_pressman/status/1457171648293924867 (accessed on 20 April 2023).
- Qiao, Han, Vivian Liu, and Lydia Chilton. 2022. Initial Images: Using Image Prompts to Improve Subject Representation in Multimodal AI Generated Art. In *Creativity and Cognition (C&C '22)*. New York: Association for Computing Machinery, pp. 15–28. [CrossRef]
- Rassin, Royi, Shayli Ravfogel, and Yoav Golberg. 2022. DALL-E-2 is Seeing Double: Flaws in Word-to-Concept Mapping in Text2Image Models. *arXiv arXiv:2210.10606*. [CrossRef]
- Ringer, Jeffrey. 2003. Faith and Language: Walter Hilton, St. Augustine, and Poststructural Semiotics. *Christianity and Literature* 53: 3–18. [CrossRef]
- Saharia, Chitwan, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily Denton, Seyed Kamyar Seyed Ghasemipour, Burcu Karagol Ayan, S. Sara Mahdavi, Rapha Gontijo Lopes, and et al. 2022. Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. *arXiv arXiv:2205.11487*. [CrossRef]
- Singler, Beth. 2020a. 'Blessed by the Algorithm': Theistic Conceptions of Artificial Intelligence in Online Discourse. *Journal of AI and Society* 35: 945–55. [CrossRef] [PubMed]
- Singler, Beth. 2020b. The AI Creation Meme: A Case Study of the New Visibility of Religion in Artificial Intelligence Discourse. *Religions* 11: 253. [CrossRef]
- Skala, Matthew. 2022. Twitter thread: Okay: Yes the Pictures Are Freaky But the Phenomenon Is Not Really as Surprising as It May First Seem [. . .]. *Twitter.com*, September 7. Available online: <https://twitter.com/mattskala/status/1567300206969982979> (accessed on 20 April 2023).
- Smith, Ethan. 2022. A Traveler's Guide to the Latent Space. Available online: <https://sweet-hall-e72.notion.site/A-Traveler-s-Guide-to-the-Latent-Space-85efba7e5e6a40e5bd3cae980f30235f> (accessed on 20 April 2023).
- Stable Diffusion. n.d. *Stablediffusionweb.com*. Available online: <https://stablediffusionweb.com/#demo> (accessed on 20 April 2023).
- Swanson, Steph Maj. 2022a. Twitter Post: To Clarify for the Press (Many Are Asking for an Explanation without Jargon): I Have Brought a Real IRL Demon to Life [. . .]. *Twitter.com*, September 8. Available online: <https://twitter.com/supercomposite/status/1567664073759678464> (accessed on 20 April 2023).
- Swanson, Steph Maj. 2022b. Twitter Thread: I Discovered This Woman, Who I Call Loab, in April [. . .]. *Twitter.com*, September 6. Available online: <https://twitter.com/supercomposite/status/1567162288087470081> (accessed on 20 April 2023).

- Tassi, Paul. 2022. Meet Loab, The AI Art-Generated Demon Currently Haunting The Internet. *Forbes*, September 7. Available online: https://www.forbes.com/sites/paultassi/2022/09/07/meet-loab-the-ai-art-generated-demon-currently-haunting-the-internet/?utm_campaign=socialflowForbesMainTwitter&utm_medium=social&utm_source=ForbesMainTwitter (accessed on 20 April 2023).
- Wang, Jianyuan, Ceyuan Yanh, Yinghao Xu, Yujun Shen, Hongdong Li, and Bolei Zhou. 2021. Improving GAN Equilibrium by Raising Spatial Awareness. *arXiv* arXiv:2112.00718. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.