Framing the Eyes:

Generative AI and the Development of Algorithmic Seeing

Ada Popowicz

14460696

MA Philosophy of the Humanities

University of Amsterdam

June 17, 2024

Supervisors:

Dr. Selin Gerlek

Dr. Stefan Niklas

Table of Content

1. Introduction

The basement is dark. The subdued murmur of other spectators and the occasional squeak of wooden chairs punctuate the silence. Although the seats offer little comfort, the basement's warmth provides a welcome respite from the biting cold of a Parisian January. Suddenly, the room is illuminated as a burst of light strikes the wall. The only other sound is the mechanical whirring of a peculiar device at the back of the room. Projected onto the wall is the familiar sight of the Gare de La Ciotat near Marseille. A beautiful black and white postcard. But wait. That is no static image. It is alive! The people there are moving, waiting, carrying bags, walking through the station! And something else is approaching… It is the train. The train! It keeps going. It is not slowing down. Is this merely a trick of camera obscura? How close is the actual train then? Why isn't it stopping? It's heading straight for us! Move! Run! Everybody! To the back of the room!

That is likely a description of an audience member's experience in the first-ever public screening of a commercial movie made by the Lumière brothers: *The Arrival of a Train at La Ciotat* in 1896. The legend has it that although the movie was not in colour but in black and white, and had no sound aside from the background whirring of rolling Cinématographe (Museum of Modern Art, n.d.; Karasek 1994), the 50-second projection of a train coming at the audience (Lumière and Lumière 1896) caused an intense reaction. People reportedly screamed and fled from the screen during the projection, which was held in the basement of Paris's Grand Café (Gregor and Patalas 1962; Karasek 1994; Eisner 1983).

On February 15th, 2024, a similar reaction of the audience's panic could be observed on Twitter. Sora, a new video-generating AI model created by OpenAI stirred reactions of both awe and fear. Prophecies of the end of cinema as we know it (Lee and White 2024), cries about loss of jobs (Chhetri 2024), the joy of tech enthusiasts and OpenAI shareholders (Metz and Mickle 2024), and fear of all those shocked by how quickly the technology was improving, was only heightened by Sam Altman, CEO of OpenAI, who on his Twitter, after announcing the arrival of Sora, encouraged Twitter users to give him prompts that he could feed to the AI software in real-time, to show the new technology's capabilities (Altman 2024). The short text-to-video outputs were rich in detail and showed a deep understanding of the real world and interactions between objects in that world, complete with the correct application of light and laws of physics.

Both of the groundbreaking technologies, Sora and Lumières' movie, were not longer than one minute and still managed to leave a lasting impression on their audiences. Both were not the first creations of their kind (Lumières' productions were preceded by many moving image experiments and Sora is not the first text-to-video generative model), but both had new features that changed the course of the history of filmmaking (for Lumière brothers, it was the presentation format—screening for a new audience in a new setting, for Sora it was the outstanding quality of the outputs that exceeded incomparably the outputs of other text-to-video models)—a "remarkable moment" in the words of Sam Altman (Altman 2024). There are, however, even more similarities

between the first-ever commercial movie shown in public and the contemporary AI creations. Both the moving image itself and the audience's reactions are highly fabricated.

The scene captured by the Lumière brothers seems natural and untouched in its realism—the passersby just going through the platform, acting as if the camera was not there. That, however, is quite surprising, since the Lumières's device was a novelty and would surely stir some reactions if not at least curious looks. Turns out that the scene is more controlled and directed than it would seem. The passersby were informed of the filming ahead and asked by the Lumière brothers to not look at the camera (Grundhauser 2016). The scene seems so natural precisely because it is taken out of actual reality. The camera replaces the view of the human eye, which, lacking the camera's imposing physical presence, would not cause any strong reactions. The human eye is invisible in the crowd—the camera is an unavoidable revelation and all it can do is replicate, in a controlled way, the human eyes' perception.

According to scholar Martin Loiperdinger, the story of the audience running out of the cinema is no more than an urban legend popularised already while the film was screening (Loiperdinger 2004). The film was shown on a significantly smaller screen than in contemporary theatres (barely over 200cm wide), it was constantly flickering and the lack of sound or colours was replaced by an abundance of film grain. It is highly improbable that even audience members not used to three-dimensional movement on a screen could confuse this film for reality (Grundhauser 2016). According to Loiperdinger, the idea of "the train nearly crashing into the audience" was merely a metaphor used by the news coverage of the time to describe the force of the new technology (Loiperdinger 2004). One might even wonder if the legend might have been popularised by the Lumière brothers themselves to promote the power of their new cinematic device.

The fabricated nature of videos created by text-to-image software is readily apparent. Every frame is synthetic, with nothing capturing actual reality. There is no need to conceal the presence of a camera, as the Lumière brothers attempted, because there is no camera to hide. Another fabrication, often overlooked, is the way technological companies, represented by figures such as Altman, use smart and powerful rhetoric to convince the public of the unbeatable power of their new products. Just as the urban legend brought lasting fame to the Lumières, the fearful buzz surrounding Sora generates attention and profit—both in terms of shareholder trust and public interest—for OpenAI. The legend of Sora, amplified by public reactions and perceptions, translates into significant profits for its creators.

New technologies, especially visual ones, have a history of inciting fear (more often than enthusiasm), as exemplified by OpenAI's new model. Photography, now an almost unnoticeable part of daily life, underwent similar processes that generative technologies are going through today. Photography, initially a groundbreaking advancement over painting, offered a new way of capturing reality. Today, AI presents an almost human-less and reality-less method of simulating the world in a synthetic manner. Interestingly, as the fear of a new technology fades and both audiences and producers become habitualised to it, an even more fascinating process occurs: our

perception and experience of the world transforms. New habits of seeing and practices emerge. It is no longer just humans shaping technologies; technologies are now reciprocally shaping humans.

This transformation in perception due to technological advancements has been explored by numerous scholars, such as Walter Benjamin (2008), Władysław Strzemiński (1947), John Berger (1972), and especially Susan Sontag (2008), who paid particular attention to photography. In the philosophy of technology, a pioneering figure in the analysis of the relationship between humans and technology was Don Ihde, who expanded the earlier ideas into a wider spectrum of optical technologies such as microscopes or telescopes and argued that optical devices transform our ways of seeing and hence the scientific ways of interpreting the world (Ihde 1999).

In 1935, Walter Benjamin, in his famous essay *The Work of Art in the Age of Mechanical Reproduction*, investigated how the effects of mechanical reproduction change how we engage with art and how the loss of physical aura alters our view of cultural artefacts (Benjamin 2008). In the work, he argued that technological advancements democratise art by making it more accessible, but simultaneously risk diminishing its unique value and presence.

In 1947, a Polish artist and theoretician, Władysław Strzemiński, finished his book *Theory Of Seeing*. There, he argued that our view of the world is constantly evolving. The ways in which we perceive it are shifting and expanding, heavily influenced by external conditions such as historical experience, and cultural and social background. Strzemiński claimed that seeing and thought are in a constant process of co-evolvement (Strzemiński 2016, 52), often aided by technological development.

In 1972, in *Ways Of Seeing*, John Berger explored the importance of artwork interpretation. In his analysis, Berger relied on the idea of "reading" images, a hermeneutic concept explored later by Ihde in terms of technology and "reading" the artefacts (Ihde 1999). Berger examined whose perspective is reflected in the world depicted within paintings and argued that art reveals the social and political contexts of its era. Furthermore, he asserted that as society and culture evolve, so do our ways of seeing.

Among these early theoreticians, the most relevant contribution to this thesis was delivered by Susan Sontag, who in her 1977 essay *On Photography* introduced the concept of "photographic seeing." She argued that the invention of the camera transformed the way humans look at the world around them. According to Sontag, the invention and widespread use of the device have not only transformed how people use it but also how they interact with and perceive their environment. She observed that experiencing a landscape through a camera lens differs from direct observation, and noted that the act of travel has become akin to consumption. Photography, she argued, has become "a way of certifying experience" (Sontag 2008, 12) of being in a particular place, often reducing the tangible, embodied experience to a mere souvenir (ibid.). According to Sontag, photography has alienated us from the real world and has become "one of the principal devices for experiencing something, for giving an appearance of participation" (ibid., 14). In her view, our perceptual frameworks and our mentality have changed—we started looking at the world "as a set of potential photographs" (ibid., 11) and suddenly, thanks to

the technology, reality exchanged its interconnectedness and continuity for manageability and opacity. Here again, technological mediation gradually shaped the perceptions and experiences of its users.

Of course, the comparisons between the public reception of photography and generative AI, especially in the context of moral, aesthetic and legal discussions is an issue already discussed by multiple scholars (Chesher and Albarrán-Torrer 2023; Manovich 2023). However, my use for such a comparison is to elucidate the development of a new type of technical seeing, caused by the arrival of a new technology: generative artificial intelligence.

In the following thesis, I aim to extend the notion of technical seeing and, using a postphenomenological lens, both in terms of theory and methodology, explore how the introduction and proliferation of visual generative technologies have created a process of "framing the eyes'' and a phenomenon of "(post-)algorithmic seeing." Through the concept of "framing the eyes," I will demonstrate how continued interaction with a specific technology induces a process of habitualisation in its users, gradually transforming their perceptual frameworks and practices. By referring to the phenomenon of "(post-)algorithmic seeing," I will explain how human perception mediated by generative technologies leads to shifts in visual and aesthetic frameworks and how AI shapes our understanding of reality.

One of the core theories in postphenomenology is that technologies and humans continuously and mutually co-shape each other. Technology plays a fundamental role in shaping human experience, perception, and practices, and even mediating our value frameworks (Kudina and Verbeek 2019). Photography, being a widely used technology, carried us through the process of acquiring "photographic seeing" (Sontag 2008). Now that we are habitualised to that way of seeing, a new one, caused by a different technology, arrives—but here, we are just at the beginning of the process of developing "algorithmic seeing." Much can be said about the ongoing process and much can be speculated about its future consequences. This research aims to offer a comprehensive analysis of how generative technologies reconfigure our perceptual frameworks, practices, and experiences, impacting not only the creation and reception of culture but also our fundamental understanding of human-technology relationships, what constitutes art, definitions of concepts such as "creativity," and the human experience in the digital age.

This study is guided by the following research question: *How do the technological stabilities introduced by visual generative AI impact users' perception, habits, and experiences and what kinds of subjects are produced through the technologically-induced process of "framing the eyes"?*

To answer this question, I will deeply engage with postphenomenological theory through the literature provided by some of the key authors in the field of the philosophy of technology, paying particular attention to seminal figures such as Don Ihde and Peter-Paul Verbeek. In the chapter "State of the Debate," I will establish the theoretical framework at the intersection of human experience and technology, discussing seminal postphenomenological concepts, and explain my postphenomenologically-informed methodology.

In the chapter “Algorithmic Seeing,” in order to explain the phenomenon of algorithmic seeing and its first occurrences, I will explore the cultural and perceptual effects of viral AI-generated content, tracing the intense period of visual generative AI development in March 2023. I will use Sora, OpenAI’s newest text-to-video model announced in February 2024, as my main case study and examine the technology in a multi-faceted way, from its capabilities and the audience’s perceptions to the introductory rhetoric used by OpenAI.

In the chapter “Framing The Eyes And Post-Algorithmic Seeing,” I will elucidate the processes guiding the framing of the eyes. I will explore what new habits of attention are forming in interaction with generative AI, how algorithmic seeing differs from naked perception, and present the importance of understanding residual algorithmic seeing, or post-algorithmic seeing. Further, I will continue my investigation by proposing new types of subjects on the production-consumption spectrum that are likely to emerge from the continued process of generative AI-human co-shaping.

Finally, in the concluding chapter, I will reiterate my key findings, emphasise the crucial role of humans as meaning-makers in the age of generative AI, and suggest how the study of this topic can be continued as time passes, technology progresses, and humans continue to be transformed by it.

2. State of the Debate

In this chapter, I will delve into the theoretical foundations and current debates surrounding the intersection of humans, the world, and technology, with a primary focus on postphenomenology. I will begin by outlining the emergence of postphenomenology from phenomenology and then explore its core concepts, such as intentionality, mediation theory, non-neutrality, mutual constitution of humans and technology, multistability, Ihde's human-tech relations and habitualisation. By analysing the key postphenomenological concepts, I aim to provide a comprehensive framework for understanding the impact of generative AI technologies on human perception and interaction with the world. This analysis will set the stage for a deeper exploration of the implications of generative AI in subsequent chapters. At the end of this chapter, I will describe the methodology used in this study, providing a clear outline of the research approach and methods employed.

2. 1. Postphenomenology

Considering the issue at hand—the process and outcomes of interactions between humans and technology—no philosophical domain offers a more suitable theoretical and methodological framework than postphenomenology. Conceptualised by Don Ihde, the postphenomenological approach examines the dynamic relationship between technologies and humans. It explores how technologies actively mediate and shape human experiences and perceptions of the world (Verbeek 2016, 7), and how humans and technologies continuously and mutually co-shape each other.

The study of postphenomenology, as the name suggests, emerged from phenomenology. According to Ihde, the technological developments of the 20th century demanded a modification of the traditional ways of thinking about technology (Ihde 2009, 8). Taking into consideration such historical contexts, Ihde attempted to introduce a more "nonsubjectivistic and interrelational" (ibid., 11) version of phenomenology, which he henceforth referred to as postphenomenology. Of course, phenomenology was groundbreaking in its own way. It claimed that through the "intentional engagement of human beings and world" (Verbeek 2005, 112), the subject and object are inextricably intertwined. Postphenomenology, however, developed the theory of intentional relationship in a twofold manner: firstly, by investigating the "fundamentally mediated character" (Rosenberger and Verbeek 2015, 12) of human-world relationship and diagnosing technologies as such mediators; secondly, by introducing the concept of co-shaping: not only are the subject and object related, but through that relation, they constitute each other (Verbeek 2005, 112). As stated by Rosenberger and Verbeek, it is that "focus on mediation and mutual constitution" that clearly distinguishes the postphenomenological approach from classical phenomenology (Rosenberger and Verbeek 2015, 12). In his study of the human-tech relationship, Ihde combined philosophy of technology with hermeneutics (Ihde 1999, 39). For him, hermeneutics, one of the "ancient" (ibid., 39) philosophical disciplines, had the potential to not only be applied to the interpretation of texts but also to the interpretation of experiences of technologies and their impact on humans.

Another important contribution of postphenomenology is its nuanced view of technology. Unlike earlier philosophies of technology, which associated technology with the loss of meaning, alienation from reality, and the end of authentic existence (Verbeek 2005, 196), postphenomenology takes a more balanced approach. It acknowledges the potential negative effects of technology while also recognizing the new ways of engaging with the world and the transformations of self-understanding that technology can facilitate, without immediately ascribing a problematized position to it (ibid., 197). In postphenomenology, technology is seen as an ambivalent, varied, and complex entity that actively mediates and co-shapes human existence.

Postphenomenology combines "philosophical analysis with empirical investigation" (Verbeek 2016, 2). Instead of forming theories and applying them to technologies, it investigates specific technologies and formulates philosophical theories based on empirical observation (ibid.). In postphenomenology, technological artefacts are not treated as mere tools or passive instruments, but rather taken seriously and analysed in their roles as "mediators of human experience and practices" (ibid.). In a crucial postphenomenological insight, the relation between humans and the world happens through things—human beings use artefacts via which they perceive the world around them, and through that mediating use, artefacts impact and actively shape human understanding and relationship with the world. In Verbeek's words, "things [...] mediate how human beings are present in their world and how the world is present to them; they shape both subjectivity and objectivity" (Verbeek 2005, 235).

Through analysing and engaging with our technologically-mediated existence postphenomenology has made significant contributions to our understanding of ourselves, our technologies and our world (Blok 2022, 3; Ihde 1999, 63). In the process, the study has developed a vocabulary of essential concepts, which will be crucial in analysing the impact of generative AI technologies on human ways of perceiving and experiencing the world.

2. 2. Intentionality

Postphenomenology is a reformulation of phenomenology. It came after it, built on top of it and has formed a separate discipline within the philosophy of technology. However, turning to and understanding some foundational phenomenological approaches and concepts can provide significant assistance in understanding the position of postphenomenology in relation to the study that preceded it and how the changes it made to phenomenological assumptions have revolutionised the philosophical approach to technology.

According to Verbeek, the most important concept of phenomenology (and one which will inform many foundational concepts of postphenomenology) is intentionality (Verbeek 2005, 108). In phenomenology, intentionality refers to the fact that consciousness is always directed *at something*. It is "the directedness of human beings toward their world" (Verbeek 2008, 13). Essentially, consciousness is directional and relational: there is no perceiver without the perceived. Intentionality is an "essential characteristic of consciousness" (Verbeek 2005, 109). As explained by Rosenberger and Verbeek, intentionality is what allowed phenomenology

to “overcome the modernistic subject-object split” (Rosenberger and Verbeek 2015, 12). However, in the case of phenomenology, the focus is strictly on the direct human-world relation. What postphenomenology does to the concept of intentionality is enriching it by adding to it an element of technological mediation. The human-world interaction is mediated by technology to form a human-*technology*-world relationship. As explained by Verbeek, such a mediating role of artefacts is what Ihde referred to as “technological intentionality” (Verbeek 2005, 114). An intriguing insight on the matter offered by Verbeek, which I shall quote in full, is as follows:

> “Postphenomenology moves beyond this modernist framework by showing that human intentionalities can not only be operative ‘through’ embodied technologies, but that in many cases ‘intentionality’ needs to be located in human-technology associations—and therefore partly in artifacts as well—without being able to entirely reduce the resulting intentionality to what was explicitly delegated to them by their designers or users.” (Verbeek 2008, 14)

In postphenomenology, it is not only human consciousness that exhibits intentionality—technologies, by dictating a specific usage, (Verbeek 2005, 114) or “providing a framework for action” (Ihde 1990, 141) form their own intentionalities “within which use patterns take dominant shape” (ibid.). A pen will dictate the speed of writing, which will be significantly different from a word processor (Verbeek 2005, 114), an AI model will allow for a different speed of production than Photoshop software would.

Of course, artefacts do not have the “determining influence” (Verbeek 2005, 115) in how they shall be used. A human user can overcome the artefact's “implicit user’s manual” (ibid.) and dictate a different course of action than the technology suggests. However, it is important to notice that when a human being uses an object, there arises a “technologically mediated intentionality,” explained by Verbeek as “a relation between human beings and the world mediated by a technological artefact” (Verbeek 2005, 116). The technology does not assume a full position of passivity. In phenomenology, the framework of intentionality is much more dynamic and interactive and, as opposed to phenomenology, it is not only human consciousness that is intentional. Technologies have such intentional power as well, which, for some scholars, further impacts attributions of other characteristics, such as imagination. As stated by Wellner, “if technologies can have intentionality that is different from human intentionality, then they can also be granted with imagination that is different from the human one” (Wellner 2022, 200).

2. 3. Mediation and Non-Neutrality

Technologies play a concrete and active role in human existence (Verbeek 2005, 11). They shape not only the relationship between humans and their environment in terms of interpretations and practices (Kudina and Verbeek 2019, 297) but also normatively: influencing entire moral frameworks (ibid., 294). The idea that technological artefacts are the mediators between the intentional relationship between humans and their world is one of the most important pillars of postphenomenological thought (Blok 2022, 3). The concept of

technological mediation focuses on how technologies shape our experiences and perceptions by mediating our interactions with the world. Technology permeates most of what we do. Artefacts occupy a mediational position (Ihde 1999, 47) between us and the world we dwell in. As noticed by Ihde, fewer and fewer actions these days are "direct or face-to-face actions, fully perceptual, actional, bodily" (ibid., 45). A classic example of a more direct action given by Ihde is picking an apple from a tree (ibid., 46). We can pick it up with our hands in a direct, i-am-my-body action (ibid., 47) but if the fruit is hanging too high, we can use a stick to reach it. Suddenly, our relationship to the apple is mediated by the simple technology of a stick. The stick "occupies [a] mediational position between the apple and me" (ibid., 47). As phrased by Verbeek, "in fulfilling their functions, artifacts do more than function—they shape a relation between human beings and their world" (Verbeek 2005, 208).

Due to technological mediation, our reach can be extended (or otherwise transformed), but the price of that magnification is a less "full" experience of the object (Ihde 1999, 47). Ihde notices that in the moment of technological mediation, our experience of the object is lessened or "reduced" (ibid.). The extent of this mediated separation can vary. For Ihde, it might involve picking up apples with a stick. In a photographic context, it could mean viewing the world through a camera rather than with our "pure" eyes. An even greater degree of separation occurs when creating a movie about a mountain using AI technologies, which generate an image that is far removed from an actual mountain, informed by representations in its database rather than a direct depiction. In a mediated situation, both what is experienced and how one experiences it changes (ibid.). In the words of Ihde, "technologies transform our experience of the objects in the world non-neutrally" (ibid.).

Non-neutrality, a concept strongly connected to mediation, negates the long-accepted view in the philosophy of science that technologies exist in a neutral and idealistic vacuum, separated from the values and norms of human society (Ihde 1999, 8). Technological artefacts are highly selective and do not simply reproduce the non-technological state—they either amplify or reduce aspects of the mediated experience (oftentimes it is that exact transformative power that makes the artefact useful in the first place), even if in a minimal way (Ihde 1999, 47).

The non-neutral transformative power of artefacts plays a crucial role in the investigation of the impact of generative technologies on our ways of seeing and AI's participation in the process of framing the eyes. Verbeek has a comment regarding technological mediation that will prove valuable for further examinations of the relationship between AI technologies and human beings in this thesis. According to him, in studying technological mediation lots of attention has been given to the artefacts themselves, but what remains largely unstudied, is the "processes of appropriating these mediations by human beings" (Verbeek 2016, 3). Simply put, in order to complete the study of technological mediation philosophy needs to pay more attention not only to the technology itself but also to the human role of giving meaning to these mediations.

2. 4. Mutual Constitution

The idea that subjects and objects are not pre-given but rather mutually shape each other is another key contribution of postphenomenology towards understanding the human-tech-world relationship. According to Verbeek, humans and technological artefacts should not be seen as two sides of an interaction—they are both products of that very interaction (Verbeek 2015, 28). Subject and object co-constitute each other in a mediated relationship. Humans and technologies continually influence and redefine each other's roles. As pointed out by Wittingslow, neither artefacts can be separated from their use, nor can humans be separated from their artefacts—"subjective perception and material objects are co-shaped; both object and user are constituted by the human-technology relation" (Wittingslow 2020, 287).

The process of co-shaping is bi-directional. Humans influence technology by use, design and modification according to the arising needs and desires, while technologies in use shape human perception, intentions and practices. This crucial interrelationality (Ihde 1999, 46) creates a causal chain driven by the introduction of new technologies, necessitating a reconsideration of existing norms and value frameworks. This results in the redefinition of social etiquette and the subsequent evolution of technological practices (Kudina and Verbeek 2019, 303). In today's technological landscape, the separate existence of humans and technologies is no longer possible—the result of constant mediation is constant mutual constitution (Verbeek 2008, 14).

2. 5. Multistability

Another important postphenomenological term coined by Don Ihde is multistability. It refers to the idea that technologies can be used and interpreted in multiple ways, depending on the context and the user. According to Ihde, "all technologies are multistable" (Ihde 1999, 49). Technologies function contextually and relativistically and there is an "essential ambiguity" (ibid.) within all of them. As stated by Wittingslow, "their interpretation and function is sensitive to (among other things) circumstance, pre-theoretic assumptions, and the individual needs of users" (Wittingslow 2020, 288). Certain stabilities are possible, but others are not.

In order to explain the concept of multistability Ihde worked with two-dimensional drawings. In his analysis, he alternated between a POV of two- and three-dimensionality, showing multiple stabilities the drawings can present depending on the context in which they are perceived. A most famous perceptual example is the so-called Necker cube. Ihde argued that if we look at the cube in the right way, we can find a different manifestation of the shape—instead of a cube, it is suddenly an insect inside a web. The drawing is open to varied interpretations but its true essence remains undetermined—"it is many things at once; it is 'stable' in multiple ways" (Verbeek 2005, 118). Ihde recognizes the same ambiguity of essence in technological artefacts. They offer separate alternations and "distinct variations" (Ihde 2009, 13) that can be fulfilled depending on the use, context or relationship within which the artefact is found.

2. 6. Human-Technology Relations

In his study, Ihde has specified and distinguished different types of human-technology relations. These relationships can vary significantly (Wittingslow 2020, 287) and with constant technological development, new relations that exceed the Ihdean categories may emerge. However, Ihde's original categorisation provides a crucial foundation for understanding the diverse roles that technological artefacts can assume. The three main types of technological relations elucidated by Ihde are mediation relations (divided into embodiment and hermeneutic relations), alterity relations and background relations.

Mediation Relations

This relation describes instances where we are not relating to the world directly but through a technological artefact (Verbeek 2005, 123). Within that type of relation, Ihde distinguishes two types: embodiment relations and hermeneutic relations.

Embodiment Relations

In this type of relationship, technology is incorporated into the user's bodily experience and forms a unity with the user (Verbeek 2015, 29). The result is an extension or enhancement of bodily capabilities which influences how individuals interact with their environment. As noticed by Verbeek, an important characteristic of artefacts functioning in embodiment relations is that they exhibit a certain transparency—"they call attention not to themselves, but to (aspects of) the world given through them" (ibid., 125). Examples include a microscope, which we look through rather than at, or a phone, which we talk through rather than to (ibid., 29).

Hermeneutic Relations

In this relationship, the users interpret reality through technology. Here, technologies influence how we perceive, interpret, and make sense of our experiences. As opposed to embodiment relations, however, the artefact here is not transparent, because it provides a representation of the world that demands human interpretation—the artefact becomes the means of perception (Verbeek 2005, 126). Examples shared by Verbeek are MRI scans allowing us to read brain activity or metal detectors giving us the ability to detect the presence of metal (Verbeek 2015, 29).

Alterity Relations

In this relationship, the technology is approached in an anthropomorphised manner (Verbeek 2005, 126). The technology forms a "quasi-other" (Ihde 1990, 101), although it can never become a genuine other. The anthropomorphisation can happen on a linguistic level: when we refer to technologies in a way we would refer to humans and project human characteristics onto technologies. Examples include humanoid robots, which mimic human appearance and behaviour, challenging our perceptions of what it means to be human and prompting reflections on the boundaries between humans and machines, or virtual companions, which simulate emotional interactions, blurring the line between human relationships and artificial constructs.

Background Relations

The last type of Ihdean human-tech relation shapes our reality without direct engagement or conscious awareness. The artefacts remain in the background but still shape our relation to reality (Verbeek 2005, 123). They are not directly experienced but still form a context of human existence (Verbeek 2015, 29). Examples can include the vibration of a smartphone on a table during a meeting, the glow of streetlights through a window at night, the faint buzz of fridges, air conditioners or fluorescent office lights.

2. 7. Habitualisation

This thesis primarily focuses on examining and analysing shifts in perceptual habits, specifically habits of seeing. Prolonged interactions between humans and technology demonstrate a transformative nature (Gerlek and Weydner-Volkmann 2022, 4). These transformations in human behaviour can occur both consciously and unconsciously. Verbeek illustrates how a microwave, by encouraging the consumption of a specific meal type, instigates a shift in lifestyle that leads to a modification in how much time people spend on meal preparation, the type of nutrition that is obtained, or even the amount of people sharing a meal—the meals are quick, lower in nutrition and often eaten solo (Verbeek 2005, 6).

Gerlek and Weydner-Volkmann in their investigation of self-tracking technologies argue that their use induces "an embodied and attentive-reflective process" (Gerlek and Weydner-Volkmann 2022, 5) thus causing a "conscious and active habitualization" (ibid.). In their paper, Gerlek and Weydner-Volkmann show that in a technologically-mediated process of gaining a habit, one can describe multiple stages of such habitualisation.

When it comes to generative AI technologies, the focus is not on reflexive self-observation but rather on a present process of viewing or production. The key interest is not in personal transformation, but rather in an observable shift in the production process that is faster, more iterative, simpler or otherwise transformed in comparison to previously available technologies. However, the individual is also transformed. The shift in human behaviour is not deliberate but rather passively acquired through interaction with technology. Gerlek and Weydner-Volkmann speak of technologies that "can become the medium of a transformation of the self-relationship" (ibid., 8). However, in the case of generative AI technologies, the focus is less on an internally directed self-relationship and more on an externally oriented perceptual transformation. It centres on how we look at, perceive, interpret and interact with what is external to us.

The development of new habits in contact with technological artefacts, especially in the context of their prolonged use has also been observed by Wellner. In her investigation of whether it is permissible to use a phone while driving, she argued that "users can develop ways to successfully split their attention between two tasks'' (Rosenberger and Verbeek 2015, 39), developing a postphenomenological notion of "multi-attention" (Wellner 2014, 6), which would make use of a cell during driving, so while experiencing two technologies at a time (ibid., 5), permissible. According to Wellner, technologies such as cell phones encourage a split of attention (thus exhibiting their technological intentionality) and humans can accommodate themselves to new technologies and "rewire

their brains'' to adapt to new emerging situations such as using a phone while driving (ibid., 6). In a human-technology-world relation, technology not only fulfils its intended use or purpose but also influences the environment in which it operates, thereby transforming human behaviour and the interaction between individuals and their surroundings (Verbeek 2005, 43) by forming habitualising strategies taken up (consciously or unconsciously) by their user.

2. 8. Methodology

The postphenomenological approach to methodology, with its distinctive stance compared to other philosophical domains, is characterised by its empirical orientation and less rigid structure (Rosenberger and Verbeek 2015, 30-32). Rather than imposing existing theories or "pre-given normative frameworks" (ibid., 31) onto technologies, it treats technological artefacts as a "concrete starting point for philosophical reflection" (ibid., 30), exploring how they (co-)shape the normative frameworks within which they exist. As the goal of this thesis is to understand how the process of framing the eyes induced by generative technologies is transforming users' experiences and perceptions in both the present and future, including the types of subjects that might emerge in the process, I will adopt a postphenomenological, empirically-informed method of examining the technology and the social discourse surrounding it.

In the next chapter, I will provide a brief historical context within which Sora has emerged. To comprehend Sora's (and visual AI technology's) social, cultural, and technological position, we must review the visual AI developments that preceded OpenAI's latest product. Despite significant advancements in generative AI throughout 2023 and 2024, the early spring season has been particularly notable for viral generative content in both years. By examining three viral AI-generated visuals from March 2023, nearly a year before Sora's announcement in February 2024, I will highlight the initial shifts in perceptual frameworks indicative of algorithmic seeing. The viral pieces of content I will examine are "Balenciaga Pope," "Harry Potter by Balenciaga," and "Will Smith Eating Spaghetti." To understand the effects of visual generative technologies on human perceptions and practices, I will focus on observing the technology from the perspective of "human beings who have a technologically mediated relation to the phenomena" (Verbeek 2016, 8). I will examine the comment sections of each piece of content in the environments where their virality was the highest—Reddit for "Balenciaga Pope," YouTube for "Harry Potter by Balenciaga," and Twitter for "Will Smith Eating Spaghetti."

In discussing my main case study, Sora, OpenAI's new text-to-video model, I will explore how it was introduced by its producer, both in terms of rhetoric and the visual outputs presented to the public. To scrutinise the corporate discourse surrounding the technology, I will examine four main sources. Firstly, I will analyse the announcement of Sora on OpenAI's blog. Secondly, I will review tweets by Sam Altman, the CEO of OpenAI, introducing Sora, along with user comments and responses showing early perceptions of Sora. Thirdly, I will describe the language used by Mira Murati, CTO of OpenAI, in a Wall Street Journal interview about the model. Finally, to understand the visual aspect of this highly visual technology, I will describe a selection of Sora's video outputs presented by OpenAI on its blog, including videos produced by human artists in collaboration with the company.

By deeply engaging with the empirical and theoretical aspects of postphenomenology, this methodology offers a robust framework for analysing the intricate interplay between generative technologies and human experience. This foundation enables a critical examination of how generative technologies and humans mutually constitute each other, the stabilities emerging from this interaction, how artificial intelligence is reshaping human perception and praxis, and the potential transformations its prolonged use might lead to in the future.

3. Algorithmic Seeing

3.1. Historical Context of Online Virality

The past few years have seen rapid advancements in artificial intelligence, marked by new AI models delivering impressive results, attention-grabbing tweets from prominent tech leaders, fierce commercial competition among tech giants, and an increasing amount of viral AI-generated content. The rise in computational power and the continuous expansion of datasets collected by large companies (Farina et al. 2024, 2) have enabled models like DALL-E, ChatGPT, Midjourney, Bard, and the latest addition, Sora, to captivate the public's attention. This has sparked widespread discussions about the value of AI-generated art, the risks of generative technologies, the definitions of creativity, job displacement, and the fundamental nature of humanity as machines progressively acquire capabilities once unique to humans. The following examples of AI-generated content and their viral online presence aptly illustrate what kinds of user perceptions developed with the introduction of each piece of artificially generated content.

Balenciaga Pope

The first piece of viral AI-generated content I would like to introduce is the image of Pope Francis wearing a luxurious white coat, commonly known as “Balenciaga Pope” (Di Placido 2023a). It was created by user @trippy_art_special[1] and originally posted on Reddit on March 24, 2023 (@trippy_art_special 2023). The popularity of the image attracted the interest of major news outlets such as Forbes (Di Placido 2023), Time (Perrigo 2023), The Verge (James 2023), Vox (Oshan 2023), and New York Times (Huang 2023). The image ignited widespread debate about the progress in generative technologies and the various stabilities AI models can assume—from dangerous deepfakes and carriers of disinformation to vessels of absurdist humour and enhancements of human creativity. While the "Balenciaga Pope'' was not the first synthetic image to go viral, it stands out for having misled a large number of people more quickly than any previous AI-generated image (Perrigo 2023; Broderick 2023). The power of the image was undeniable and it marked a “significant cultural moment” (Elias and Razik 2023). As noticed by Perrigo, “history may regard the Balenciaga Pope as the first truly viral misinformation event fueled by deepfake technology, and a sign of worse to come'' (Perrigo 2023).

The discussion on Reddit, where the original image was posted, reveals a lot about public perceptions of the image. Users expressed their shock and concern about the fact that so many people believed that the image was real (@nathanstolen 2023), “despite knowing about AI imagery” (@monsantobreath 2023a). Others admitted that the image had fooled them until they found out through another source that it was fake, even though after longer observation they could indeed notice the strange inaccuracies in the image (@luminous_lead 2023). It became clear, with the exception of some users’ opinions (@nolaconnor 2023), that the image is not indistinguishable

[1] For clarity, in this thesis, online usernames are prefixed with an “@” symbol, both in the main text and in the references. However, if users have provided their full names on their social media accounts, the standard referencing style will be employed.

from reality, but rather well adapted to the fleeting and casual glance of a regular social media user (@ArmchairReditor 2023).

The discussion took an interesting turn when the users started discussing why the image achieved such a high level of believability. According to one user, the image is "just benign enough to be plausibly real" (@supervegeta101 2023). The Catholic church is known for its wealth and it certainly could afford expensive designer clothes. One user argued that the Pope wearing a parka "like some celeb" (@monsantobreath 2023b) is an absurd idea and that users should be more sceptical about such absurdity. Others argued that puffer coats are not reserved for celebrities and it could be very believable that Pope Francis would be seen in such a garment. One user had a particularly important contribution. They noticed that even though the Pope might usually not be seen in a designer puffer coat, "in the picture which is usually evidence, he is" (@Crafty_Enthusiasm_99 2023a). They emphasised that realistic images usually hold a level of believability that is now suddenly (and dangerously) achieved by generative AI. In response to that, another user asked: "do you mean that people now should no longer believe that any picture they see is real? Because that is not going to happen" (@DPSonly 2023).

One of the more interesting outcomes of the image that fooled so many users was the emergence of "guides" aimed at educating the public on how to spot AI-generated images. An example of such an article is "How to Spot an AI-Generated Image Like the 'Balenciaga Pope'" published in Time Magazine (Perrigo 2023). Similar to other articles (Di Placido 2023a; Huang 2023) and discussions on Reddit, the author explains the aspects that indicate the synthetic nature of the image. Perrigo emphasises the importance of scrutinising details, explaining that while AI image generators are informed by their datasets about the appearance of individual objects (such as the Pope or a coat), they often struggle to reproduce the laws of physics that govern interactions between objects, making elements like gravity and shadows problematic. As Perrigo points out, "it is in these often peripheral parts of an image that humans are intuitively able to spot inconsistencies that AI can't" (Perrigo 2023). Di Placido lists the synthetic giveaways in the "Balenciaga Pope" image: "the Pope's ear is smudgy, his glasses melt into the shadow across his face, and his hand, clutching a coffee cup, is warped" (Di Placido 2023a). Additionally, his skin texture "looks waxy" (ibid.), which relates to Vincent's observation of the hyperreality of AI images. Vincent notes that AI images exhibit an aesthetic "defined by perfect lighting and glossy surfaces, by dramatic poses and saturated colors" (Vincent 2023). He argues that such qualities don't interfere much with the believability of the images, as we are already accustomed to over-stylized and exaggerated depictions of celebrities (ibid.). AI simply analyses and amplifies the trends it identifies in its dataset.

Guides like Perrigo's essentially offer instructions on how to interpret this new type of imagery, inviting readers to train (or reframe) their perception. However, as Perrigo notes, technological advancements continuously make generated results more indistinguishable from reality. For instance, he mentions how "the newest version of Midjourney can generate realistic-looking human hands, removing what was perhaps the easiest way to identify an AI image" (Perrigo 2023). He acknowledges that his advice may soon become obsolete as technology progresses (ibid.). Indeed, just a year later, in January 2024, the hosts of the popular tech podcast Hard Fork

attempted to use techniques such as examining human ears or glasses to guess whether a face was human or AI-generated (Hard Fork 2024). After several tries, it became evident that distinguishing real and fake images was almost impossible, even for industry experts well-versed in the capabilities of generative AI.

Harry Potter Balenciaga

The second piece of viral content I will present is a video titled "Harry Potter by Balenciaga," originally published on YouTube on March 15, 2023, by the user @demonflyingfox (@demonflyingfox 2023a). The video creatively merges characters from the Harry Potter series with high-fashion elements from the Balenciaga brand. The creator used Midjourney to generate the AI images, ElevenLabs to produce the character voices, and D-ID to animate the result (Panamaned 2023). Notably, "Harry Potter by Balenciaga" is a second piece of viral content that utilises the brand of Balenciaga as a generative variable. Although the photo of Pope Francis gained more media recognition, Jennings argues that the Harry Potter video is a more intriguing case study (Jennings 2023). I concur, as the most fascinating aspect of the video is that no moving footage was directly generated by a text-to-video model. Instead, @demonflyingfox fashioned an illusion of a video using static AI-generated images and sounds that captured the iconic Harry Potter characters and the distinctive aesthetics of Balenciaga (ibid.). By combining different AI tools and their functions in a new context, the creator produced a novel composite stability.

The video has proven how much the AI tools have lowered the entry to cultural production by allowing the author to create a “2 million dollar ad for probably less than 10 bucks” (@user-lm6yk4hh5z 2023). The interesting and relatively easy manner of production “has spawned a host of other popular films interpreted through the Balenciaga lens, including Star Wars, Lord of the Rings and Breaking Bad” (@Danziger 2023). One can find tutorials and process instructions both on YouTube and in popular media (PromptJungle 2023; Jennings 2023) that guide potential creators on how to reproduce a video in the style of “Harry Potter by Balenciaga.”

As noticed by a Reddit user, the video is a great example of style transfer—a technique of generative AI where the style of one image, brand, author or other cultural artefact is applied to the content of another. According to the user, “AI is notoriously good at doing this, perhaps even better than humans” (@Crafty_Entusiasm_99 2023b). Some comments express interest in a remix culture displayed by the video: “goddamn Balenciaga Dumbledore looks stylish as hell. Now I want a steampunk, gothic, matrix-ish kind of Harry Potter” (@GaaraSama83 2023). The familiar reflected under a new light has an undeniable allure in the era of ambient, unchallenging culture (Chayka 2024). Interestingly, the “Harry Potter by Balenciaga” video, which relied heavily on style transfer, through the various tutorials and reproductions, itself became a “style” that could be “transferred” to other cultural combinations.

The main concept that YouTube commenters highlighted under the video is the inexplicable and addictive uncanniness it exhibits. Examples of this sentiment include: “got no idea why i watched this disturbed clip for almost 10 times. I see no reason” (@ElectronicHouseFlash 2023), “those nods are too creepy and I can't stop

watching them" (@asdloller 2023), "there's something addicting about this it feels like some sort of bottomless abyss calling out to me" (@atheneite 2023), "I can't stop watching it." (@theseawillrise 2023), and "there's something extremely sinister yet bizarrely hilarious about this version of Harry Potter. I'd watch the hell out of it" (@casualblogger3810 2023). Users expressed astonishment at the creative use of multiple AI technologies and anticipated a further leap from AI images to AI videos.

As observed by Jennings, the "Harry Potter by Balenciaga" video wasn't merely proof that AI can fool the masses. Rather, it touches upon the potential expansion of human creativity, the lowered entry barriers to production, and raises the question of "just how long we, as a society, have before AI-powered video becomes most of what we think of as visual entertainment" (Jennings 2023).

Will Smith Eating Spaghetti

Not even two weeks after the publication of the video of "Harry Potter by Balenciaga", expectations of YouTube commenters materialised. A piece of AI-generated video content went viral. Originally published on Reddit on March 28, 2023 (@chaindrop 2023), it gained wider virality on Twitter. Made using Modelscope Text2Video generator (Di Placido 2023b), the video featured actor Will Smith eating spaghetti. However, the video was far from the current realism of Sora. As described by Di Placido, "the clip depicts a hideously malformed Smith shovelling handfuls of spaghetti into his mouth, with a terrifying gusto" (ibid.). The reactions, a mix of amazement and horror, are plentiful in the Twitter comments under the video. The sense of uncanniness is mixed with something more unsettling, as the meal merges with Smith's face and the model shows a clear lack of understanding of reality. While some users still considered the effects to be amazing (@TTEcclesBrown 2023), or at least "both magnificent and horrific" (@0xDesigner 2023), others found it "quite disturbing" (@checkthreetimes 2023) or an "atrocity" (@EddyVGG 2023), with some even expressing that they "will never sleep again" (@Beersnake_21 2023).

Initially, Di Placido noted that the only potential stability of generative video technology involved "horror, and a touch of nausea" (Di Placido 2023b). He also observed that early AI-generated images often featured laughable imperfections such as "melted faces, wonky, missing limbs, and extra fingers" (ibid.). However, month after month, the technology improved, and the "disturbing" imperfections evolved into hyper-realism. Another Twitter user shared Di Placido's comparison to early AI-image technology and speculated that by 2025 we might see "amazing text-to-video" (@Anthis 2023). Just a few months later, Sora delivered on that promise—already in 2024.

The "Will Smith Eating Spaghetti" video demonstrated a keen interest in the advancement of generative technology toward video content. Its virality had a lasting impact, and following the release of Sora, many users compared the original video with new outputs from Sora (Brownlee 2024; John 2024), often expressing amazement and disbelief at the rapid pace of technological progress. The Will Smith video did not deceive anyone; the synthetic nature of the video was obvious, and viewers did not need to adjust their perceptual frameworks. Today, identifying synthetic elements is not as straightforward. However, despite that impressive

development, the current generative technology is not perfect. Educated users have developed habits of detecting indicative signs such as melted deformations, waxy skin textures, strange hand gestures, and other significant details. Unfortunately, less informed users, who are unaware of the existence and functioning of generative technologies, remain vulnerable to disinformation and are slower to develop algorithmic literacy and algorithmic seeing.

Professor Leaver, from Perth's Curtin University, noted that our algorithmic literacy does not evolve “at the speed of the tools that can create material" (Elias and Razik 2023). As technology continues to progress, it will become increasingly challenging for users to develop the refined methods of seeing needed to identify algorithmic productions. The believability of photographic images is no longer assured. As Leaver expressed, “photographic-level imagery and truth just need to be completely disassociated" (ibid.). This technological evolution demands a transformation in both perceptions and habits of seeing.

Each discussed example of viral generated content highlights different aspects of the evolving human-technology relationship. A key observation is that photographic believability is no longer a reliable perceptual assumption. Generative AI, approaching an unprecedented level of realism, has been able to deceive the public with images like “Balenciaga Pope” already over a year ago. It is important to note that the image of the Pope successfully fooled many people not only due to its technological sophistication but also for two additional reasons. Firstly, the image was “low-stakes” (Di Placido 2023). Unlike depictions of political figures or war events, the Pope’s fashion choices do not constitute a “world-changing event” (ibid.). Users could easily scroll past it without further scrutiny because the image held little significance beyond a fleeting moment of humour or curiosity. Secondly, in many instances of its digital presentation, the fake photo of Pope Francis was juxtaposed with other unusual but real photos, such as one of him “holding a microphone and signing a Lamborghini” (ibid.). This created an interesting phenomenon—“real photos imbued the fake image with a veneer of authenticity” (ibid.). In a visual reality guided by algorithmic output, the combination of fake and real images results in perceived authenticity.

The “Harry Potter by Balenciaga” video, however, succeeded not by pretending to be realistic but by showcasing new possibilities for cultural reality and using the technology in a completely different way, thus establishing new stabilities. With prompt-based generative AI software, staples of popular culture such as Harry Potter or Balenciaga can be reduced to mere variables or style transfers that, when wielded skillfully, produce new and often exciting forms of entertainment that challenge our existing aesthetic frameworks. The shift in technological capabilities, moving from the laughable horrors of “Will Smith Eating Spaghetti” to the uncanny hyperrealism of Sora, and the prompt-based nature of the process, creates space for a new form of entertainment based on style transfer and remixes of existing culture.

3. 2. Case study: Sora

In a setting already shaped by content like “Balenciaga Pope,” “Harry Potter by Balenciaga,” and “Will Smith Eating Spaghetti,” OpenAI has unveiled its latest breakthrough: Sora, a text-to-video model capable of producing

hyper-realistic moving images. Announced on February 15, 2024, this technology has yet to be made publicly available but has already sparked intense debate and evoked a mix of fear and amazement among those who have seen its outputs. Since the software has only been announced and not released, this analysis will primarily focus on how the technology was introduced and discussed, the examples of its outputs that were shown, and how users perceived the technology.

3. 2. 1. Corporate Discourse around Sora

One of the most intriguing avenues of postphenomenological research is examining how a company that created a novel technological product introduces it to the public and attempts to shape social expectations, perceptions, and future stabilities of the technology. OpenAI introduced Sora through several channels: by publishing the announcement on its website (OpenAI 2024a), through tweets by Sam Altman (Altman 2024), OpenAI's CEO, and via an appearance by Mira Murati, OpenAI's CTO, in a Wall Street Journal YouTube video covering the video model (Murati 2024).

Sociotechnical Imaginaries and Website Announcement

Nothing illustrates OpenAI's deliberate rhetoric better than the name of its new text-to-video model: "Sora." The OpenAI research team chose this name, derived from the Japanese word for "sky," to evoke "the idea of limitless creative potential" (Metz 2024). Additionally, in a market primarily oriented towards the West, the use of a Japanese word may convey notions of precision, technological advancement, and beauty, which are often stereotypically associated with Japan. The intentional use of language does not, however, end with the name. The blog post from OpenAI announcing Sora (OpenAI 2024a) features many sentences that reveal the narrative that the company wants to create around the tool.

The focus of the software is expressed in concepts such as realism, complexity, detail and accuracy. According to OpenAI, Sora "is able to generate complex scenes with multiple characters, specific types of motion, and accurate details of the subject and background" (ibid.) and can "create realistic and imaginative scenes" (ibid.). However, it is also underlined that the software is human-oriented. It is supposed to "help people solve problems" (ibid.) and be able to not only understand human language on a "deep level" (ibid.), but also understand and reproduce "vibrant emotions" (ibid.). From a technical point of view, what also seems to be stressed is the ease with which the technology can be used: "solely from text instructions" (ibid.)—the barrier to entry could not be lower. Finally, OpenAI reveals what Sam Altman frequently states as the company's ultimate goal: creating artificial general intelligence (AGI). As stated by the company, "Sora serves as a foundation for models that can understand and simulate the real world, a capability we believe will be an important milestone for achieving AGI" (ibid.). AGI refers to a highly advanced form of synthetic intelligence capable of performing any intellectual task that a human can, demonstrating flexibility and understanding across a wide range of activities. That goal is not secret. When one wants to apply for a job in OpenAI, the very first point of introduction to the company is a statement about its AGI focus: "we are committed to building safe, beneficial AGI that will have a massive positive impact on humanity's future" (OpenAI 2024b).

The news coverage following the release was generally favourable towards the new software. A journalist from Wired noted that Sora not only fulfils user prompts but does so "in a way that shows an emergent grasp of cinematic grammar" (Levy 2024). Levy also explained that OpenAI considers its new product to be especially "distinguished by its striking photorealism" (ibid.). Furthermore, Bill Peebles, an OpenAI researcher interviewed by Levy, expressed his wonder about the fact that Sora generated its outputs with sophisticated and narrative-based use of camera angles, cuts and timing—and more importantly (and more magically) did so on its own, without explicit instructions from the engineers (ibid.).

The main type of critique that one could expect from the media coverage of Sora is that of potential risks and dangers of such a technology being released to the public—the capacity for deepfakes, disinformation and general chaos seems plausible, especially in the year of US elections. However, throughout the release post, OpenAI makes a point of creating guardrails for the technology's use, promising the application of multiple safety methods and consultation with "domain experts in areas like misinformation, hateful content, and bias" (OpenAI 2024a). That of course is a remarkably smart move for two reasons. Firstly, the media and the public cannot criticise the company too harshly for something that it has openly admitted. Secondly, underlining the importance of controlling such software strongly suggests just how powerful the technology can be. And few things sell better than power.

One thing is clear: OpenAI is trying to control the narrative of Sora's perception and use. Such a strategy is not new and has already been studied in 2009 by Jasanoff and Kim under the name of "sociotechnical imaginaries" (Jasanoff and Kim 2009). They define the phenomenon as "collectively imagined forms of social life and social order reflected in the design and fulfilment of nation-specific scientific and/or technological projects" (ibid., 122). Although their project focused on how governmental narratives dictate the reception of nuclear technology, their conclusions can be easily applied to any narrator and any technological product—including AI companies and AI products. Jasanoff and Kim elucidate that when it comes to the early days of a new technology it is less about what the technology actually does, but rather what users (directed by a larger body of power) think of its behaviour and how that influences the human-technology-world relationship. Their research indicates that governmental narrative profoundly shapes society's ability to envision the future impacts of new technologies. They contend that imagining potential futures is crucial in influencing social and political landscapes. Imaginaries function as interpretative frameworks, helping individuals understand their experiences and directing their actions and decisions.

In the case of the algorithmic technology discussed in this study, one theoretical development of sociotechnical imaginary is worth mentioning: Bucher's concept of "algorithmic imaginary" (Bucher 2017). Bucher's key project was to reverse Jasanoff and Kim's top-down order of how technological understanding is proliferated and give more agency to the users, looking at how "people imagine, perceive and experience algorithms" (ibid., 31) and "the productive and affective power that these imaginings have" (ibid., 41). Algorithmic imaginaries are

influenced by cultural, political, and historical contexts, as well as societal narratives about algorithms. They encapsulate collective attitudes, beliefs, and values regarding algorithms and highlight expectations about their impacts on individuals, communities, and institutions. In the modern era of hyper-connectivity, users create meaning communally, making it much more challenging to enforce one-directional technological propaganda.

Once again, revisiting the time of the introduction of photography as a new technology offers a useful comparison. As noted by Sontag, in the early days of personal cameras, “Kodak put signs at the entrances of many towns listing what to photograph. Signs marked the places in national parks where visitors should stand with their cameras” (Sontag 2008, 54). Such a clear direction of “what to photograph” seems strange, even ridiculous to us, modern camera users. We don’t need such explicit direction anymore. Why? We already acquired photographic seeing. We speak the visual language fluently and photographically consume reality in ways that early Kodak executives couldn’t dream of (even though we switched to different products than those offered by their company). Photography is now normalised and so is photographic seeing. AI-generated technology, however—is just at the beginning of the process. And it is in the interest of its producers to control how it will be normalised. Although perhaps not explicitly, their consultants know that “by means of declaration and status functions we create social reality” (Coeckelbergh 2015, 180). To put all the mentioned findings in postphenomenological terms, the public perception of the technology in a specific context will influence the stabilities the technology will assume. Sometimes, the development of perception can happen even before the technology is actually widely used, which is why producers of new technologies have a vested interest in influencing user perceptions as early as possible. Just like Kodak put out signs of what to photograph, companies like AI will try to influence the habits that the users will acquire as they interact with a technology like Sora.

The ease of use promised by OpenAI, with videos created "solely from text instructions" (OpenAI 2024a), follows a narrative inherited from earlier technologies. As Sontag observed, users appreciate frictionless, easy, "invisible technologies" (Sontag 2008, 17). Manufacturers assure users that capturing photos requires neither skill nor expertise, claiming that the "machine is all-knowing, and responds to the slightest pressure of the will" (ibid.). For cars, it is ignition; for guns, it is the trigger; for photography, it is the shutter; and for generative AI, it is the prompt. The technology has a new stability: an irresistible and addictive "fantasy-machine" (ibid.).

Fabricated Sublime and Altman’s Tweets

OpenAI’s CEO, Sam Altman, has an undeniable understanding of the power of storytelling. Altman is as skilled a narrator as he is a businessman, adeptly using rhetoric to steer media, investors', and public attention in the company's favour. On Twitter, he commands a following of nearly three million users, with each post drawing significant attention. The release of Sora was no different. His post announcing Sora (Altman 2024a) gained over six million views. Praising his team’s achievements with superlatives like "incredible" and "amazing," Altman described the appearance of Sora as a "remarkable moment" (ibid.). He invited Twitter users to suggest ideas for prompts in the comments, confidently encouraging them to not "hold back on the detail or difficulty" (Altman 2024a). Almost a thousand comments answered his encouragement (in addition to the announcement post from

just 2 minutes earlier, which gained almost two thousand comments). Users eagerly challenged the software with their ideas, and Altman enthusiastically responded. After seeing Sora's outputs, commenters immediately reflected on the impact of Sora on the movie industry: "this is going to disrupt the current VFX industry for sure. Producers will hug you" (Deshpande 2024), "film industry (in its current form) is so fucked" (DePew 2024), "bearish Pixar and Disney" (Halford 2024). Others worried about the wider social implications and the pace of AI development: "probably gonna need to start figuring out UBI or some form of it as a society sooner rather than later lol" (@EddyScamArtist 2024), "AI goes too fast. Let us absorb for a sec" (Hassid 2024). However, the majority of comments seemed to express a mix of astonishment and fear: "this is unreal" (Ford-Monroe 2024), "wow. This is going to change everything" (@TheAvatarMovies 2024), "for the first time ever, we can literally recreate our own nightmares" (@Olney1Ben 2024), "what's that magic?" (@ent4151 2024), "the future is starting to look like the future" (Dhesi 2024).

The experience of fearful-pleasurable awe is a concept well-known in philosophy as "sublime." The concept, although vast, can be briefly described as a specific type of aesthetic experience that evokes a "thrill or shudder of perverse pleasure, mixing fear and delight" (Morley 2021). In Kant's early conceptualization, the sublime arises when humans encounter phenomena of overwhelming size, depth, power, or transcendence. This confrontation leads to a state of awe that is both terrifying and pleasurable, as individuals grapple with the limits of their perception and understanding (Burnham, n.d.). This blend of terror and delight is central to the sublime experience, highlighting the tension between human vulnerability and the grandeur of nature or art.

The mixed reactions to Sora, which exemplify a sublime experience, are an advantage to OpenAI. Both fear and excitement can be turned into profit. Through the narratives he puts forward, Altman is trying to establish stabilities of the technology with the help of fabricated sublime—of course focusing on the stabilities that will be most beneficial to OpenAI. Generative AI models, in their inherent ambiguity, can be seen as helpful or as dangerous. But most importantly—they are powerful. That is the stability that Altman wants to promote. Admittedly, the technology we are experiencing is astonishing, even to the industry experts, but the sublime is artfully fabricated to create the "legend of Sora." And the early adopters, who often grew up on science-fiction are a perfect audience for such storytelling. Altman's company directs both the production of AI products and the production of the sublime surrounding them. In this process, users' perceptions are framed, and control over the imaginary of the new technology is developed—along with control over the technology's stabilities and users' habits.

However, users are not powerless. Altman's attempts at subliminal storytelling are often met with criticism. For instance, in a recent tweet regarding Sora, Altman claimed that "movies are going to become video games and video games are going to become something unimaginably better" (Altman 2024b). While many users accept Altman's narratives unquestioningly, others are more critical. One user commented, "can't movies just stay movies. I like those" (@KettlebellDan 2024), to which another cynically replied, "Nah, movies are 'content' now" (Fryant 2024). A particularly insightful critique came from a user who noted, "Sam can't make money unless he

keeps pitching fancy-sounding tech ideas to big-money donors. He doesn't really care if [...] those ideas totally make sense... or if anyone even actually wants them" (@RationalEye 2024). Altman's goal is not factual accuracy, but rather generating provocative and attention-grabbing narratives. Of course, users can refute the imaginaries dictated by the tech companies, but new imaginaries will emerge nevertheless—possibly from more socially informed sources rather than corporate narratives. However, one aspect will remain true: technology, in its non-neutrally mediating nature, with the help of its producers or without it, will continue to influence human perception.

Aesthetic Quality Dictation and Murati's Interview

The final introduction of Sora that I want to explore is Mira Murati's interview with The Wall Street Journal. In the interview, Murati engages in what I refer to as "aesthetic quality dictation," a phenomenon specific to visual, creative, or artistic technologies. A very specific type of medium is promised, and particular qualities are highlighted as important. In describing Sora, Murati uses a precise set of words to convey the technology's characteristics. The technology is not only easy to use—"just based on a prompt" (Murati 2024)—but is also described as "hyper-realistic," "highly detailed," "beautiful," and "smooth" (ibid.). While the company has thus far focused on capability, it now aims for optimisation—making the technology more steerable, controllable, and accurate (ibid.).

At first glance, Murati's language seems technically oriented, as her role as CTO would dictate. It is natural for an AI manufacturer like OpenAI to view technology in technical terms. However, as a company positioning itself as an aid to artists, its non-artistic rhetoric cannot be overlooked. Indeed, later in the interview, Murati assures in an artist-aiding spirit that "AI tools will extend our creativity, knowledge, collective imagination, [and] ability to do anything" (ibid.). Although many technologies before Sora were also spoken of in technical, rather than artistic terms by their manufacturers, no Kodak camera producer claimed that its purpose was to enhance humanity and provide the "ability to do anything" (ibid.)

In the interview, whether consciously or not, Murati highlights what is considered a desirable experience of artistic quality under Sora. The interview serves as a guide in perception, instructing not only on how to perceive the product itself but also on how to view the world, reality, art, and human-technological relations in a "Sora-like" manner. This perception-guiding aesthetic value dictation, though not directed by a specific corporate agent, was also observed by Sontag in the context of photography: "in teaching us a new visual code, photographs alter and enlarge our notions of what is worth looking at and what we have a right to observe" (Sontag 2008, 8). OpenAI actively engages in such aesthetic assessment. While traditional yet elusive artistic values like "beauty" are mentioned, aspects such as "smoothness" and "hyper-realism" do not necessarily apply to the broader goals of all imaginable genres of art. However, OpenAI positions these characteristics as indicators of the company's success and markers of artistic quality.

Murati does not speak of Sora's ability to create aesthetically challenging images or pose existential questions to its audiences through art—the language is simpler and optimised. This terminology matters. The language used to describe new technology "encodes a culturally specific viewpoint" (Bones et al. 2021, 27). It teaches viewers what to look for and what to find aesthetically satisfying—it guides their perception. Language serves as a tool for imbuing technology with meaning (Coeckelbergh 2015, 180). While only humans can be the source of that meaning (ibid.), those in power can gently nudge the masses in a particular direction through fabricated sublimity and control over imaginaries. By iterative storytelling and sustained narratives, new user habits and broader human-technology relations become, at least partially, shaped by the company that manufactured the technological product. Technologies have their materiality, but it is human perceptions of these technologies that establish the reality of the stabilities associated with Sora (Coeckelbergh 2023, 10).

3. 2. 2. Visuals Made with Sora

As evidenced by how OpenAI describes Sora, they aim to present it as part of a new generation of creative tools. Like the other AI-generated outputs previously discussed, Sora went viral, but its virality was not based on a single perfected output, but rather on the subliminal influence of the model as a whole. However, Sora is ultimately a product—a product that will soon be for sale. The company needs to avoid portraying it purely as a product, as that could invite uncomfortable questions about the sources of scraped data, the cost and conditions of human labour, sources of investment, environmental impact, and more. Emphasising creativity and humanity is a much safer strategy. It is preferable for users to perceive the technology as a means of enhancing and supporting human creativity, rather than as a questionable tool built on illegally obtained data and displacing human jobs.

As a visual creative tool, Sora's verbal introductions were accompanied by a variety of video outputs: a flock of paper planes, an elegant woman walking down a Tokyo street, mammoths traversing a snowy landscape, vintage-looking shots of astronauts in wool-knitted helmets, a drone shot of a sunny cliff, a papercraft coral reef, pirate ships battling in a coffee cup, and historical footage of a Gold Rush-era American city (OpenAI 2024a). At first glance, the videos display impressive variety. However, all the videos share the visual features characteristic of the viral media discussed in the previous section ("Balenciaga Pope," "Harry Potter by Balenciaga," "Will Smith Eating Spaghetti") or those described as desirable by OpenAI representatives.

The first set of features, unsurprisingly, includes imperfections—although, as with "Balenciaga Pope," one must look closely to find them. In the context of a video, investigating these imperfections involves more friction. Similar to the other outputs, Sora's creations were not perfect in terms of replicating physics or reality, although they showed significant improvement over technologies from the previous year. Users could use their newly trained eyes to spot inconsistencies in details, such as mismatched hand movements, elements melting into one another, or objects disappearing and reappearing without clear causality (Heaven 2024). However, detecting such imperfections in a moving image is more challenging than in a static one. One must zoom in, take time to examine, and even stop frame-by-frame to study the shot both in motion and still. OpenAI, perhaps consciously,

created additional friction for detailed examination—in their blog release, one cannot pause the videos (OpenAI 2024a), leaving screenshotting as the only tool for investigators.

It is worth noting, however, that Sora acknowledges the faults of its model. In a separate blog section about "room for improvement" (ibid.), some videos clearly display the shortcomings mentioned. When one clicks on a tiny plus next to the image caption, a description of the image's weakness appears—though one "weakness" caption cuts off mid-sentence: "simulating complex interactions between objects and multiple characters is often" (ibid.), under a video of a grandma celebrating her birthday with her family. This could be an honest web-design mistake, but considering the attention a company like OpenAI gives to its online image, it is notable that the one part "wrong" with the website is a description of a shortcoming of their newly released product. Despite this miscaptioned image, other imperfections are correctly critiqued, highlighting "inaccurate physical modeling," "unnatural object morphing," or "physically implausible motion."

Cinematic Smoothness

Another common feature in all of OpenAI's videos is their tempo. Each video moves its contents smoothly and slowly across the frame. In cinema, slow motion is often used to dramatise a scene, enhance emotional effect, or create a sense of anticipation. Such effects would undoubtedly benefit the public perception of Sora. This slow, smooth storytelling choice likely stems from the type of prompt used to produce the images—"cinematic." The "cinematic" prompt has long been a staple in the prompt engineer's toolkit, well before Sora's introduction. As of June 2024, this prompt had generated over 18.5 million results on the Midjourney Discord server (Midjourney 2024), demonstrating its perceived effectiveness in achieving impressive text-to-image outputs. As explained by Trillo, "cinematic" is a desirable feature for AI filmmakers (Trillo 2024). However, should "cinematic" be a characteristic applied to all cultural outputs? Overusing special effects like slow motion can diminish their communicative power over time, causing them to lose their impact. Consequently, filmmakers might need to reduce their reliance on such effects or replace them with new storytelling techniques.

Smoothness in Sora's videos is observable not only in motion but also in texture. Even with objects that seem correct at first glance, the strange "glossiness" of the surfaces reveals itself to those who know what to look for (similar glossiness can be seen in both "Balenciaga Pope" and "Harry Potter by Balenciaga"). At this stage of algorithmic seeing, informed users notice the glossiness and "know what to look for"—while less algorithmically literate audiences can be easily fooled and remain unsuspicious about the texture of objects. As the technology becomes more widely applied and the perceptual frameworks develop, such glossiness will either disappear due to technical optimization or become normalised and unnoticeable. In both cases, the strangeness of texture will become invisible as user perception adjusts to the new technology. Features we currently associate strictly with synthetic videos and images will become normalised as we grow accustomed to a reality increasingly shaped by AI cultural production.

Art-Washing and Sameness of Outputs

A little over a month after Sora's announcement, on March 25, 2024, OpenAI released a new set of outputs generated with the model. Unlike the initial releases, these results were not cherry-picked by the company but created by artists, designers, and filmmakers who were given early access to the new tool. According to OpenAI, the purpose of this collaboration was to "learn how Sora might aid in the creative process" and "help creatives bring ideas to reality" (OpenAI 2024c)—in postphenomenological terms, to explore the stabilities of the technology that OpenAI aims to promote.

The videos, or rather short films, created by artists are slightly longer than the original creations taken straight from the model in OpenAI's previous post, where the videos lasted only a few seconds. Here, the creators had the freedom to use external editing tools. Another significant difference is that these creations are more narrative-driven and include audio (which does not come from Sora, as the tool currently only has visual capabilities). These videos are not mere snippets of content but representations of how artists can engage with the tool. In the video "air head" by shykids (OpenAI 2024d), the software performs impressively in applying physics. The balloon on the main character's head reflects surrounding objects, and the car mirror reflects the balloon, although the balloon's reflection in the water is less successful. Technical mistakes in the background can be observed, but the ongoing narrative diverts attention away from them.

Despite the presence of narrative, critics have pointed out the vapidity of these films (Ranquist 2024), particularly given the complexity and multi-disciplinary nature of cinema. It is "a depressing artform to see AI being used on" (ibid.). Ranquist argues that choices guided by life experience are what make cinema great, and AI lacks such life experience. Of course, it can be counter-argued that humans (the meaning-makers) still make decisions about what to generate and which generated options to select and combine. The debate about values guiding art and creativity (Bueno et al. 2024) continues in an infinite loop.

The critical opinions come from cinephiles, individuals deeply committed to film, who make video essays about it or at least voice their strong opinions in comment sections. As Ranquist points out, most regular viewers will not care about the qualities of "real movies" (Ranquist 2024)—the experiences of specific users will shape their perception of the technology and the stabilities they can envision for it. According to Ranquist, for the majority of regular users, the potential of the technology will be more exciting than the artistic value it can produce (Ranquist 2024).

However, artists who have collaborated with AI defend its artistic position, and OpenAI eagerly includes these endorsements in its blog post (OpenAI 2024c). Shykids note that "as great as Sora is at generating things that appear real - what excites us is its ability to make things that are totally surreal" (OpenAI 2024c). Musician August Kamp claims, "Being able to build and iterate on cinematic visuals this intuitively has opened up categorically new lanes of artistry to me... I truly cannot wait to see what other forms of storytelling will come into reach with the future of these tools" (OpenAI 2024c). Josephine Miller adds, "The ability to rapidly

conceptualize at such a high level of quality is not only challenging my creative process but also helping me evolve in storytelling" (OpenAI 2024c). It is difficult to find a word of critique about the tool from the artists who have worked with it so far. Each of them highlights new stabilities of the technology they noticed during their experimentation—easy iteration, rapid conceptualisation, and the development of new storytelling methods—all very positive perceptions of the technology's abilities.

Director Paul Trillo is at the forefront of creators who seem to have successfully incorporated Sora into their workflow. After releasing the short film made for OpenAI, he explained that "Sora is at its most powerful when you're not replicating the old but bringing to life new and impossible ideas we would have otherwise never had the opportunity to see," and stated that "working with Sora is the first time I've felt unchained as a filmmaker" (OpenAI 2024c). When faced with the tool, he was shocked, floored, confused, and unsettled, as he did not expect the technology to be so capable (Newton and Roose 2024). His goal was to make his film "cinematic" (a quality I highlighted earlier), and to "unstick" it from the "AI-looking video" aesthetic (ibid.).

The criticism from cinephiles like Ranquist aligns with accusations directed at OpenAI regarding their introduction of Sora. The company has been accused of art-washing (Newton and Roose 2024), a practice of using art to gloss over or distract from controversial issues (Pritchard 2023)—a clear attempt to portray the software as "artistic." OpenAI's current practices are perceived as leveraging artists to showcase and test AI tools, ostensibly highlighting their creative potential while potentially positioning these tools to replace costly human labour—tools that may have been trained on the work of the same artists now at risk of being displaced (Newton and Roose 2024). Similar to OpenAI, Trillo asserts that the tool is not intended to replace humans but to serve as "a much better alternative to stock footage B-roll" (ibid.). He emphasises that traditional tools and methods remain essential for achieving nuance and detail in films.

Despite this stance, Trillo recently released "the first official commissioned music video" using Sora (Trillo 2024). The video used only the text-to-video model and received mixed opinions. Twitter fans and followers of Trillo mostly expressed positive reactions (ibid.), but commenters on Reddit and YouTube, less associated with Trillo's fanbase, were more critical. The perceptions varied significantly between the two audiences. The extent of criticism is evident in the most upvoted comments under the Reddit post about the music video (@adesigne 2024): "It's just the same fucking SORA video that literally everyone is producing: walking through incoherent visuals that are superficially rational but in reality make absolutely no sense" (@no-one-4845 2024), "Wow, I can't believe I'm saying this but I'm already bored of Sora and it's not even out yet" (@midnightmiragemusic 2024), "The manic infinite zoom is getting stale" (@UnequalBull 2024), "Wow, that's extremely uncreative" (@Kanute3333 2024). The main accusation seems to be that Sora-based creations, even those made by professional creators, are boringly similar, devoid of artistic value, and reveal the tool as more of a gimmick than a creatively transformative device. After the initial awe (mixed with terror), users are starting to express a certain level of disillusionment with the technology and dissatisfaction with its repetitive aesthetic. Perceptual frameworks are transforming, but not in a direction that could be seen as beneficial for Sora's producers.

The developments discussed in this chapter indicate a significant shift in user perceptions, driven by the capabilities of generative AI technologies and the narratives constructed around them. The exploration of viral AI-generated content, such as "Balenciaga Pope," "Harry Potter by Balenciaga," and "Will Smith Eating Spaghetti," reveals the dual nature of this shift. On one hand, these technologies have the power to awe and captivate audiences with their hyper-realism and creative possibilities. On the other hand, they raise critical concerns about authenticity, the erosion of photographic believability, and the potential for widespread misinformation.

The emergence of Sora, OpenAI's advanced text-to-video model, exemplifies these dynamics. Through corporate rhetoric and user perceptions, it becomes clear that while there is excitement about the unprecedented ease of creating realistic videos from simple text prompts, there is also a growing awareness of the repetitive and somewhat superficial aesthetic that often accompanies these outputs. OpenAI's efforts to control the narrative surrounding Sora's perception highlight its interest in shaping how this technology will be normalised. The public perception of the technology in a specific context will ultimately influence the stabilities the technology will assume.

Although users initially exhibited a sense of awe and fascination with the technology, many are gradually becoming more disappointed and disillusioned with its capabilities. OpenAI continues to influence sociotechnical imaginaries, develop fabricated sublime, and engage in aesthetic quality dictation to establish powerful stabilities that sell well and benefit the company, but the users do not always follow OpenAI's storytelling.

This blend of user fascination and scepticism underscores the importance of human agency and the role of meaning-making in critically engaging with these technologies. As we navigate the evolving terrain of algorithmic seeing, it is crucial to maintain a reflective stance, recognizing both the transformative potential and the inherent limitations of AI in shaping our visual reality. Standing at the cusp of algorithmic seeing, it is evident that this is merely the beginning of a longer, more complex process of framing the eyes, one that will continuously reshape our visual and conceptual engagement with the world.

4. Framing the Eyes and Post-Algorithmic Seeing

As observed by Kudina and Verbeek, “corporate discourse co-shapes users’ perception of their technologies” (Kudina and Verbeek 2019, 300) and OpenAI’s narratives about Sora fit that observation perfectly. The company actively positions its product on the technological-creative stage according to what is beneficial to its (profit-driven) goals. That includes influencing users’ perception of the technology, including what makes its outputs beautiful or desirable and what kinds of stabilities it should assume. Neither the mediating technology, nor the company that produced it has a determining influence on public perceptions, and many users continue to steer away from the way of seeing promoted by OpenAI. However, the ongoing debate over how technologies like Sora should be perceived as just the beginning of the process of framing the eyes.

In this chapter, I will deepen my investigation into not only the current perception of generative AI and the phenomenon of algorithmic seeing, but also extend my study to understand the process of framing the eyes and its future implications. In the first three subchapters, I will cover various aspects guiding this process, including new habits of attention that develop with AI exposure and the differences between naked perception and technical seeing. Crucially, using previously discussed examples, I will highlight the importance of focusing not only on algorithmic seeing but also on residual or post-algorithmic seeing in our understanding of the relationship between humans and generative technologies. In the final subchapter, I will explore and speculate on the types of subjects that can emerge from the interaction between human beings and generative AI technologies, considering both producers and consumers in the process of human-technology mutual constitution.

4. 1. New Habits of Attention

After being exposed to a few eerily realistic generative AI creations, such as “Balenciaga Pope,” many started looking at the existing visual content with a certain level of suspicion. People don’t like being “fooled” by AI and after gradual exposure to more and more AI-generated content, they started the process of “learning-to-see” to make algorithmic content visible to them. The idea that technical vision is a skill to be learned, in Ihde’s writing, was already clear to Galileo, who “was well aware of the need to teach telescopic vision” (Ihde 1999, 178) and became so infatuated with the telescopic way of seeing, that he proclaimed that the technologically-mediated perceptual framework was superior to the “ordinary vision” (ibid.). It is a crucial postphenomenological insight that, as we shape new technologies, they mutually shape us. Many technologies gradually habitualise us to see the way they “see” and the process is usually not immediate. People working in a specific industry, for example, textile, gain “expert seeing” that allows them to see details in the threading that are invisible to ordinary viewers. Such expert seeing might be technologically mediated or not, but it is acquired upon repeated exposure to a particular context. In today's digital landscape, younger people, more used to and aware of technological mediation and details of the functioning of AI technologies, are much less likely to mistake synthetic content for the real thing. They possess a form of digital literacy akin to expert seeing. Conversely, older generations,

especially the so-called "boomers," who are more technologically disconnected, are more prone to mistaking generated content for authentic photographs. This issue is apparent in the widespread virality of AI images on Facebook, driven by the enthusiastic responses of older users to the generated content (Tangermann 2024).

One reason members of less technologically connected communities find it more difficult to recognise synthetic content is that they have been habitualised into a photographic way of seeing, which includes photographic believability. In human judgement, photography tends to automatically carry the value of truth. As observed by Magnus, “when we see features in a photograph, our default rule of evidence is to believe them” (Magnus 2023, 3). A similar fact was observed by Sontag: “a photograph passes for incontrovertible proof that a given thing happened. The picture may distort; but there is always a presumption that something exists, or did exist, which is like what’s in the picture” (Sontag 2008, 10). Ihde adds that “by the 1890’s photographs had become the standard recorders of objective scientific truth” (Ihde 1999, 179). Painting or other forms of artistic depiction are not expected to be truthful in a way that a photograph is. It is more difficult to imagine that the contents of a photograph are not real. However, recent viral AI creations, such as “Balenciaga Pope,” have challenged the assumption of photographic believability.

According to Chesher and Albarrán-Torres, AI-generated content “can never make the truth claim that photographs often still retain” (Chesher and Albarrán-Torres 2023, 67). But in the era of hyper-realistic AI outputs, no photography is free from suspicion. An interesting consideration here is the case of computational photography. In many mobile devices, companies like Apple or Samsung have “enhanced” the camera by adding default AI settings (Chayka 2022) which, for example, take multiple images and then stitch together the most interesting parts to form a “perfect image” (perfect from the perspective of algorithmic evaluation of what is real and aesthetically desirable). Suddenly, photography is also algorithmic. Perhaps no image can be “fully” real anymore—if it ever was, as Sontag argues that photography, like other technologies, is non-neutral (Sontag 2008, 11).

Frequent exposure to a specific technology transforms the perceptual framework. In the initial stage of algorithmic seeing, that transformation is tied to visual suspicion—people don’t want to be “fooled” by AI. As discussed in the previous chapter, some of the strategies include looking for “hyper-reality,” strange “glossiness” or paying more attention to anatomical elements, such as human hands or ears, whose mechanics and structure are difficult for AI models to accurately reproduce. The process taking place in users’ ways of seeing can be described as “forming a habit of attention” (Gerlek and Weydner-Volkmann 2022, 13). Users pay more attention to visual content, focusing on different elements than they would if another technology was used. One must be aware of the phenomenon to perceive it. Looking is one thing, but seeing is another. As noted by Strzemiński, “we only notice the phenomena to which our attention is directed. Our thoughts seem to preemptively pose questions to which our vision seeks answers” (Strzemiński 2016, 54).

With every technological era, society faces new tasks that necessitate observing emerging phenomena reflective of the time's essence. To truly grasp them, one must adapt their method of observation. Extended interaction with a specific technology forces modifications in ways of seeing, practices, and frameworks of behaviour (Gerlek and Weydner-Volkmann 2022, 4). As habitualisation continues, the transformative nature of the technology becomes more naturalised and thus invisible, at which point the process of framing the eyes with that specific technology can be considered complete. Ihde, using the example of getting used to new glasses, describes the phenomenon: "once learned, the 'irreality' either is diminished or virtually disappears as the instrument is properly 'embodied' into the new (now normalised) experience" (Ihde 1999, 156).

While generative technologies are still at relatively early stages of development, some transformations in habits of attention can already be observed. To avoid mistaking synthetic content for reality, users now scrutinise areas that did not require particular attention in the era of photographic seeing. By examining hands, ears, glasses, suspicious glossiness, hyper-reality, or deviations from the laws of physics, users hope to distinguish the real from the generated. This is an attempt to develop expert technical seeing that exposes algorithmic influence. Additionally, it is becoming clear that an important new habit is to let go of the assumption of photographic believability—a particularly challenging shift for older generations who grew up with photographic seeing and are deeply habitualised to it. As AI technology progresses and reproduces reality with increasing accuracy, the assumption that a photographic-looking image is a recording of objective truth is one that cannot safely be carried into the future.

4. 2. Naked Perception versus Technological Mediation

If technologies possess the transformative power of co-shaping and mediating human experiences, then visual technologies have the capacity to transform our perceptual frameworks. For Ihde, mediation is "indissolubly linked with a transformation of perception" (Verbeek 2005, 130). He argues that there is a fundamental difference between "naked perception" (Ihde 1999, 157) and perception through artefacts, and this difference often encompasses the non-neutrality of the technology (Verbeek 2005, 130). What Ihde calls "naked perception" (Ihde 1999, 157) has also been termed "natural vision" (Sontag 2008, 44), "direct seeing" (Magnus 2023, 2), or can be referred to as organic or biological seeing. Considering the postphenomenological angle of this thesis I will follow Ihde's "naked perception" in referring to the way of seeing that does not rely on technological artefacts.

Like Ihde, I speak of unmediated vision in terms of vision that lacks technological intervention, where perception is not mediated by technology (Verbeek 2005, 125). Human perception is always mediated in some way because we access the world only through interpretation, which already imposes an interpretative filter on reality. Technology is merely one of many possible mediators, but the absence of technological intervention does not imply that a "preinterpretive access to reality" (ibid.) will occur.

On one side, we have artefact-free "naked perception," and on the other, we have technical seeing. This includes "expert seeing," "telescopic vision" (Ihde 1999, 178), photographic seeing, and at the centre of my investigation, algorithmic seeing. Any imaging technology can be argued to have transformed perceptual frameworks by imposing a new filter on reality. As Ihde observes, "It makes a great difference if one can view the heavenly bodies only with one's naked eye or with the device of the telescope" (ibid., 43). The mediation relation (whether embodiment or hermeneutic) between the user and the technology significantly impacts human perception.

However, algorithmic seeing is distinct. While still a form of technical seeing, it functions differently from its predecessors. Its relationship with reality differs from other imaging technologies, which magnify, constrain, or otherwise frame actual reality. AI is synthetic. Its results are not straightforwardly related to reality. It can, of course, be argued that paintings were also not straightforwardly related to reality. Even when figurative, they revealed the sensitivity and subjectivity of the artist. Photographs, according to Magnus, have been consistently mechanical (Magnus 2023, 1)—a perspective many photographers would dispute, arguing for individual sensitivity in their works. But AI's detachment goes further. There is a difference between seeing a person or a picture of a person and seeing an AI-generated image of a person (ibid., 2).

Any tool and its prolonged use contribute to a shift in the way of seeing, and every artefact will have a different effect depending on the stability it assumes in a specific context. Certain aspects of the original "naked perception" will remain unchanged, while others will be strengthened, weakened, or otherwise transformed (Verbeek 2005, 131). However, some form of transformation through the human-technology relationship is unavoidable. Chesher and Albarrán-Torres aptly compare generative technological outputs to simulacra to explain AI's relationship with reality. They claim that generated media reproduce what was never seen, "a resemblance without a referent" (Chesher and Albarrán-Torres 2023, 67). Their contribution is worth quoting in full:

> "Rather than shaping light with a lens into analogue images, autolographic algorithms do not work with images. They work with digital information and calculations of probabilities in dataspaces. Its subjects are not in the visual fields of everyday life but in calculations of data patterns in massive image archives. Autolography fixes its images not through chemistry or CCDs, but through programmed random diffusion procedures in neural networks" (ibid.)

The entirely novel process of AI image production is unlike any preceding technology. Its foundation is script and data-based, and the resulting output is a visualised calculation. What we perceive are numbers dressed in pixels that, through advanced computation, pass as realistic depictions of actual reality. The resulting transformations in ways of seeing are profound and long-lasting. Since algorithms don't work with images but with "digital information and calculations of probabilities in dataspaces" (ibid.), they are fundamentally different from other imaging technologies. Consequently, algorithmic seeing can be fundamentally dissimilar to previous forms of

technical seeing. AI's novel relationship to reality has the capacity to transform human perception in ways we have not experienced before and cannot yet fully understand.

4. 3. Relics of Residual Seeing

A transformation in the way of seeing might be understood as something temporary, occurring only when the mediating artefact is in use. However, many technologies leave a lasting impact that extends beyond their direct use. The influence of certain technologies transforms our way of understanding the world in such a fundamental way that it is impossible to imagine how pre-technological assumptions could be seen as acceptable. Ihde offers an acute observation on this phenomenon. He recounts how, while observing the moon through a telescope, he wondered how "the ancients could ever have thought the Moon to have a pure, mirror-like, featureless surface" (Ihde 1999, 157). For him, it was simply impossible to see it that way. According to Ihde, his perception became post-telescopic, and the Moon appeared differently to him after acquiring telescopic knowledge. The eyeball "contains" within itself the "'residue' of the telescopic" (ibid.). The Moon looked different to Ihde, whether viewed through a mediating telescope or with the naked eye (ibid.). In Ihde's words:

> "This is, in short, a lifeworld accretion which follows an irreversible direction. While there are different Gestalts for naked and mediated perceptions, there is also an interaction and overlap which through familiar embodiment shapes the contemporary texture of the lifeworld" (ibid.).

As Ihde observes, "technologies as perception-transforming devices not only magnify (and reduce) referent phenomena, but often radically change parameters either barely noted, or not noted at all" (ibid., 163). Algorithmic seeing involves not only the embodiment relation of seeing "through" an algorithmic interface but also a long-term hermeneutic relation, where algorithmic seeing reconfigures the "sorting" of reality even when no algorithmic artefact is in use. This new way of seeing challenges the agreed-upon interpretation of the world on a societal level. Sontag, writing about photographic seeing, proclaimed: "what is true of photographs is true of the world seen photographically" (Sontag 2008, 65). Photographs have transcended their role as mere recordings of reality and have become "the norm for the way things appear to us, thereby changing the very idea of reality, and of realism" (Sontag 2008, 70). With Ihde's and Sontag's insights in mind, it becomes clear that the notion of photographic believability is carried over from the now-normalised photographic seeing. In the era of AI, it is a relic, a form of residual seeing that does not match the reality of the new algorithmic technology. Surrounded by generative tools, users need to adjust their (post-)photographic perceptual framework and let go of the habit of attention that assumes photographic believability.

As I mentioned earlier, Ihde pointed out that "by the 1890's photographs had become the standard recorders of objective scientific truth" (Ihde 1999, 179). Considering that photography was invented already in the 1820s, that means that the process of photographs becoming emblems of objective scientific truth was gradual. Assuming that the 2020s are to generative AI what the 1820s were to photography, we have a very limited capacity to

understand not only what exactly algorithmic seeing will become, but even less what effects post-algorithmic seeing will have. If photography as a mediating technology can have the lasting and powerful impact that Sontag describes then surely the same can be expected of algorithmic technologies.

How is the world viewed algorithmically in ways that we can already observe? We have already seen several examples. Our perceptual and aesthetic frameworks are being restructured by prompting and corporate discourses. Users are increasingly adopting expert seeing to guard against being "fooled" by synthetic imagery. New parameters, considered potential giveaways of AI, are being identified more frequently. Cultural icons, such as Harry Potter or Balenciaga, along with various artistic styles and techniques, are now perceived differently—as mere cultural variables that can be assembled into a prompting composition for content generation.

An important aspect of what I argue to be algorithmic seeing lies in the residual aspect of interacting with AI technologies. What is or will become the residue of the algorithmic way of seeing? What phenomena have we started seeing only after being exposed to and working with algorithms, which have now become normalised and incorporated into our algorithm-free, not-actually-mediated reality? What aspects of experienced reality became noticeable after the introduction of algorithms, and which ones disappeared? These questions beg for graspable examples of residual algorithmic seeing, but the answers cannot yet be provided. We can speculate that, as technologies are perception-transforming devices (Ihde 1999, 163), our relationship to reality itself will change significantly with the use of AI. Unfortunately, most questions about residual algorithmic seeing, or post-algorithmic seeing, are unanswerable with the current level of AI development and human habits of seeing. As Chayka observes, "it takes decades, if not centuries, to determine just how a technology has influenced cultural forms [...] only when a new tool has become unremarkable can its effects be judged" (Chayka 2024, 291).

What we can do at the moment and in the near future is to observe the early hints of the development of new perceptual frameworks and, by looking back at the history of technological-human relations, try to infer how the process of framing the eyes mediated by a specific technology might continue. Only through such intentional observation can we hope to retain our agency over how we shape technologies and understand how they shape us.

4. 4. New Co-Produced Subjects

Generative AI technology has not yet been widely adopted. The current reactions come from those who are already interested in the technology (even if only out of fear), rather than from those who will experience its fully normalised production and reception. These early adopters are not the final subjects that will be co-produced with the technology through the process of framing the eyes and (post-)algorithmic seeing. Since the subject holds a central position in postphenomenological study, in the following section, I will propose what kinds of new subjects might emerge as AI generative technologies transition from new-to-the-world to normalised.

When considering creators, I will focus on how the different stabilities of AI technologies will reshape the workflows of artists, expand the spectrum of their expressive possibilities, redefine the aesthetic frameworks under which they work, transform their imagination, and alter their perceptive relationship with technology and the world. For consumers, I will analyse the tension between AI-assisted personalization of media and the homogenised flattening of culture, exploring how new habits of seeing might expand or limit our experience of culture and the world.

Oftentimes, algorithms (and hence algorithmic ways of seeing) are associated with a sense of loss: loss of human elements, human authenticity, or simply human jobs. However, in the postphenomenological spirit, it is essential to take a nuanced look at technology and avoid the technological romanticism that characterised phenomenology. In my proposal of the types of subjects created through the process of co-shaping with algorithmic technologies, I will strive to avoid immediately ascribing them to a "problematized position" (Verbeek 2005, 197). Guided by Verbeek's assertion that "technological artifacts indeed close off some possibilities by the way they mediate experience, but they also open up new ones" (ibid., 203), I will explore ways in which algorithmic seeing might enrich the human experience.

4. 4. 1. Creators and Producers

The introduction of new technologies transforms not only the ways creators work but also their understanding of "the aesthetic values underlying their works" (Farina et al. 2024, 3). This has been evident with the invention of oil painting, photography, Photoshop, and generative AI. Creativity, although considered fundamental to human intelligence (ibid., 6), is a very elusive, multifactorial, and complex topic (Boden 2004; Bueno et al. 2024), which I will not attempt to define in this thesis. It is, however, important to emphasise the postphenomenological aspect of creativity and imagination.

As Wellner observes, creativity and imagination transform over time due to their technological co-shaping (Wellner 2022, 190). They operate very differently in the prehistoric age, Renaissance, after the introduction of the camera, or in the digital age (ibid.; Strzemiński 2016). The mutual constitution and perception between humans and technology heavily impact users' abilities to incorporate technology into their practices, thus influencing the position of technology in a specific society and the definition of concepts such as "creativity." Each era is defined by its tools, which open up new "experiences and possibilities" for "artists, viewers, and society as a whole" (Wellner 2022, 200). Imaginative faculties are not ahistorical but flexible and exist in "co-shaping and co-constituting relationships with our technologies" (ibid., 191). As Verbeek states, "designing technology is designing human beings" (Verbeek 2015, 28)—including their imaginative frameworks and creative practices.

"Cinematic" Prompting and Cultural Capital

As observed by Sontag, "photography in Europe was largely guided by notions of the picturesque [...], the important [...], and the beautiful" (Sontag 2008, 52). An important notion, or more explicitly a prompt, that guides current AI production is that of a "cinematic" quality. New visual technologies, such as generative AI, shift the

desirable aesthetic from a more static “picturesque” (scenic, harmonious, serene) to a dynamic “cinematic” (vivid, detailed, dramatic)—a shift clearly influenced by the development and widespread use of moving image technologies over the past decades. The use of specific technologies has mediated the experience of the world (and what is beautiful in it) in a non-neutral way.

Interestingly, from a producer's standpoint, the ability to explicitly dictate characteristics such as “cinematic” directly to the technological device generating the image is a notable development in the artistic process, made possible only with prompt-driven text-to-image models. Fifty years ago, photographers could not simply apply the prompt of “picturesque” to an image. The process of imbuing the work with such qualities was more implicit and depended on the artist’s technical skills to achieve a specific visual goal.

In contrast to analogue photography, generative technology's foundation is script-based. It is connected not to ineffable artistic skills but to explicit prompts. A dramatic change introduced by prompting is that what matters is not technical skill but rather knowledge of possibilities. People interested in art who have not gained an artisan skill set, but through interaction with art have developed an artistic way of seeing, now have an avenue of expression for that vision. Prompts have allowed for a particularly curious development: their use reduces art tools, techniques, and even artists into mere variables that can be optimised for a desired result through prompt engineering. Algorithmic art-making is akin to shopping in a department store filled with visual taxonomies and technical guides on how to best use their evocative power to obtain the most impressive AI-produced results.

While it may initially appear that creating quality art has become more accessible, a crucial factor that often distinguishes successful AI art is rich cultural capital. To generate something using AI, prompters need an understanding of what looks good—a developed perceptual framework gained through previous experiences, whether technologically mediated or not. Online guides and lists of “best prompts” can only go so far in informing the user about what will look good when generated. Thus, amateurs, in competition with experienced artists, often find themselves at a disadvantage. Artists have built up rich cultural capital. Through their practice, their perception has been framed to prioritise certain aesthetic frameworks over others and notice things that regular users of a technology (such as a camera or a paintbrush) would not. Their work transforms their perception and guides the development of specific technological stabilities. Similarly, those working with technologies like Sora will rely on their built-up cultural capital, carried over from other experiences. Over time, in a technologically co-shaped process, new perceptions, informed by experiences mediated by new generative tools, will emerge.

Certain communities favour specific aesthetics, and AI caters to some better than others. For fans of fantasy, science fiction, and video games, the frameworks of quality differ from those of a student of conceptual art. After all, “what is art differs according to historical and cultural context” (Coeckelbergh 2023, 4). The appeal of conceptual art lies not in its aesthetic but in its visual communication of an underlying concept through a metaphor or a joke—both of which AI is, so far, not very good at. Fans of Greg Rutkowski, a famous digital artist, might be pleased with AI-generated outputs prompted with his name (Heikkilä 2022), while fans of surrealist

Magritte might be disappointed with the depth of the AI-generated artwork they see on the screen. The processes underlying the works of both artists are very different—especially on the conceptual level.

If prompting becomes a leading artistic practice, more implicit or ineffable forms of art (such as surrealist or conceptual art) might be more difficult to create. They require more than a textual description of what should appear in the image. Digital fantasy artworks, such as those of Rutkowski, as beautiful as they might be, rely more on technical skill and less on the use of metaphors and thus will be easier to reproduce by an AI model. As a result, artists like Rutkowski might be more easily replaced, as their art relies less on the deep meaning-making skills that are uniquely human.

Art that embraces humour or ineffability, indicative of high-level meaning-making skills, might thrive as artists seek creative avenues where they can successfully compete with machines. However, this type of art could also diminish if the costs of creating conceptual art become significantly higher, or if the general audience, influenced by the pervasive presence of AI-generated imagery and the socio-technical imaginary promoted by AI companies (which champion characteristics such as "cinematic" as hallmarks of true art), begins to prefer "simpler" AI art over conceptual art. This preference might arise because AI art is perceived as "cinematic" or possesses other traits deemed indicative of artistic value, which will evolve as the process of framing the eyes continues.

Future art production will depend on material circumstances and evolving definitions of what constitutes art. This will not hinge on any mystical essence of AI outputs, but rather on how human perceptions are shaped and conditioned to evaluate aesthetic desirability. Currently, it is the "cinematic" quality that seems to gain an important position in framing the eyes.

The Power of Defaults

Algorithmic video creation introduces a completely different production method compared to the pre-algorithmic era. Complex tools come with many options for users to choose from. However, the more complex the tool, the more interesting it is to look at the default settings set by the manufacturer. Defaults impact the use of technology—as noticed by Verbeek, "the default settings of copy machines and printers help to determine how many double-sided prints will be made" (Verbeek 2015, 29). For example, Midjourney automatically generates images of conventionally attractive people (Nyce 2023). This is likely because in large datasets of faces, when a "generic human" is generated, the faces get averaged and tend to be judged as attractive by viewers (Iyengar et al. 2015). While skilled prompt engineering can override this default, its existence still influences creators and the visual landscape they can produce.

A particularly tangible example comes from researchers who noticed that "when ChatGPT is set to low temperature and asked to pick a random number, the odds are higher than chance that it will pick forty-two" (ibid.). Why? Socio-historical curiosity. Historically, programmers, who more often than not are fans of science-fiction, like to reference the popular book *The Hitchhiker's Guide to the Galaxy* by Douglas Adams. In there, a supercomputer, after 7.5 million years of thought, "reveals that the answer to the 'Great Question of Life, the

Universe and Everything' is 'forty-two'" (Delahaye 2020). For the geek culture, the number became something between an inside joke and a fixation (ibid.). Although this one is a rather harmless example of a default, it does show the fact that the phenomenon is real and it can non-neutrally impact both the creators and the generated content.

Real issues, whether positive or negative, arise when technology manufacturers use their agency to set defaults intentionally. OpenAI filters portrayable reality, forbidding porn, gore, and public figures (Murati 2024). Through default settings, companies can dictate the contents and qualities of art forms, thus transforming user perception. Certain styles or characteristics can be promoted, while others are suppressed, significantly influencing the produced visual landscape. This "act of censorship" (Chesher and Albarrán-Torres 2023, 64) allows companies to control portrayable reality, impacting the creative freedom of artists and influencing the stabilities of the technology. While most measures taken by companies are claimed to be guided by safety concerns (OpenAI 2024a), one has to wonder what other measures underpin the production of "algorithmic culture" (Chayka 2024, 44).

If the boundaries of portrayable reality are shaped by the technology used, the workflows of AI artists (or the experiences of regular users) will be non-neutrally influenced, and the users themselves will be co-shaped by the technology. Initially, these effects might seem insignificant. For example, if public figures cannot be generated, political satire would have to rely on pre-AI tools. Similarly, if it becomes impossible to prompt for gore or nudity, certain forms of art (and harmful imagery) might diminish. However, the effects can be more profound: with AI predominantly rendering attractive people by default, viewers' perceptions of reality, particularly regarding average human appearance, may become skewed. This can exacerbate body image issues and influence personal care habits to align with evolving beauty standards.

The notion that a single company could control the stabilities of AI technologies is problematic, especially as they become more widespread and users become habitualised to relying on them. As the evolution of social media has shown, with the growth of technology's popularity and user habitualisation to its daily use, alternatives may become unaffordable or unavailable. Control over the technology remains with the owning company, allowing it to dynamically adjust the "guardrails" based on its objectives. As demonstrated by social media, this is not always beneficial to humans.

Stabilities and Human Action

Through mediation and co-shaping, technologies can have a transformative effect on human imagination and technical abilities, subsequently influencing habits and practices. This transformative relationship is bidirectional: "we imagine new technologies, and they draw new horizons for our imagination" (Wellner 2022, 191). While technological artefacts might draw new horizons, the data-based nature of generative technology could also prove limiting—restricting the spectrum of possibilities to what is already in the dataset and reducing the emergence of "the new" (although whether any artwork can be truly new is another contested topic that falls

outside the scope of this thesis). AI-assisted art production offers new possibilities but also introduces new limitations compared to previous technologies. As Ihde states, technologies do not simply “reproduce the non-technological state—they either amplify or reduce aspects of the mediated experience” (Ihde 1999, 8). This observation is crucial to remember when producers of AI tools try to present the technology as merely extending human capacities. Different stabilities of the technology will limit one aspect while expanding our reach in another (Ihde 1999, 47).

For instance, while fast iterative features allow artists to create multiple versions of a visualisation of the same prompt, maintaining character consistency has always been one of the biggest challenges for producers of text-to-image models, even referred to by some as “one of the holy grails of generative AI storytelling” (Morrison 2024). The same difficulty can be expected to arise in the context of text-to-video technologies. In traditional filmmaking, the camera simply follows an actor, making character consistency not only successful but inevitable without editing. This challenge is already evident in AI filmmaking. The most featured video created by artists in collaboration with OpenAI using Sora is "air head" by shykids (OpenAI 2024d). The project showcases a narrative of the life of a man who has a yellow balloon instead of a head. That choice is very clever and guided by the stabilities offered by the technology. It allows the audience to connect with the story told from a human perspective while omitting the challenges of not only generating passable human faces but also character consistency.

The balloon-headed character we see on the screen is a tangible result of human creativity being co-shaped by technology. The filmmakers had to work within the limitations of the technology yet still produced impressive visual results. The design of human-tech interactions, including its restrictions, impacts “not only the design of technological objects [...] but also the design of the human subjects who interact with these objects” (Verbeek 2015, 28). The important aspect is not what the tool does, but how the tool’s functioning impacts what the user does. Transformation occurs through changes in human action. The inability of the technology to ensure character consistency influenced the creative decisions made by the filmmakers. Human creativity is relational (Blok 2022, 16), and new technologies like AI create new stabilities to which humans respond. If the tool shykids were working with had different limitations, the human course of action and resulting visual would also be different.

As Jarow observes, “building new tools is a way of building new mental capabilities” (Jarow 2023). Although the interaction is guided by the human subject, who not only gives the prompt but also evaluates “the quality and creativity of the generated work” (Farina et al. 2024, 4), the technology in use (and its design) transforms the actions and perceptions of the human subjects and shapes the patterns of use (Ihde 1990, 141). This is an excellent example of “technological intentionality” (Verbeek 2005, 114). A noticeable aspect of technological participation in intentionality is generative serendipity. The translation between language (prompt) and image is never straightforward. One can describe an image with as much detail as possible, yet the imagined visual might vary for every listener. A similar effect occurs in text-to-image generation. Often, “from the point of view of the

creator, [..] AI may do something unexpected and surprising" (Coeckelbergh 2023, 6), providing an unintended output. When this happens, "the tool, the technology, comes to the foreground" (ibid.), formulating a response that is irreducible to the creator's original intention (Verbeek 2008, 14). In that moment, Ihde's alterity relation materialises, and the technology can be seen as a "quasi-other" (Ihde 1990, 101).

This serendipity was also visible with previous visual technologies. Paint mixed with a new medium might have shown unexpected properties. Developing film in unusual conditions might have given a novel character to the photograph. An unintentional combination of buttons might have activated a Photoshop function the user did not expect. However, previous technological serendipities did not lead to anthropomorphising the tools. AI, with its opaque inner workings and advanced outputs, carries a level of mystery with its serendipitous results that prompts users to (often subconsciously) describe the technology in terms of "seeing," "understanding," or "creating"—anthropomorphic terms that might have even appeared in this thesis. The nature of generative AI challenges the perception of the boundary between human and machine, and, especially with the rhetorical strategies of AI producers, gets closer than any preceding technology to becoming not only a "quasi-other" (Ihde 1990, 101) in an alterity relationship, but even being perceived as a "genuine other" (Verbeek 2005, 126)—a notion with highly problematic consequences for human creators and humanity at large.

4. 4. 2. Audience and Consumers

Super-Personalisation and Shared Reality

Personalised algorithmic feeds have become a modern standard. From Instagram to Netflix, companies strive to tailor content to users' preferences. Abundant content must be filtered, so the selection of songs, images, or movies is personalised. The computational power of generative AI models introduces a new development: algorithmically-driven culture could target individuals not only by selecting movies but also by adjusting the content or format of movies themselves, based on an individual analysis of preferences. This possibility has significant consequences..

One primary reason why online platforms personalise our experience is their ad-based business model. By gathering data and personalising content, they assure advertisers that their messages reach the right audiences. The effect of this could be multiplied with the iterative powers of generative AI. Ads embedded in movies could, through the combined power of personalisation, enticing narratives, and AI capabilities, create a reality where it becomes difficult to distinguish between advertisements and genuine content—a challenge already present with online influencers (Sanders and Schneier 2024). While viewers might develop ways to reduce the effectiveness of such advertising, similar to how internet users have learned to skim over browser ads (Shreya 2023), subtle incorporation of advertisements into generative productions could hinder the development of such skills. Initially, users might experience suspicion and detachment as they try to distinguish art from sales pitches, but with the

right methodologies, even the most perverse advertising strategies could become normalised in the eyes of the user.

Another danger of super-personalisation is the fragmentation of shared reality. Algorithms already tend to serve us information that aligns with our existing views and then push it to extremes, resulting in social and political polarisation. The goal is not to expand users' understanding of the world or their imagination, but to find more efficient ways to keep them engaged and paying for subscriptions. Algorithms personalise our interpretation of the world, catering to our existing views rather than challenging them. As this process advances, the sense of a shared perception of reality among individuals diminishes. Each type of content or technological output experienced by a user shapes their perspective in a specific manner, which can be reinforced or altered by subsequent experiences. The perceptual frameworks that develop through interaction with algorithmic culture will be as diverse as the content to which users are exposed.

Flattening of Culture

However, super-personalisation is not the only potential path for cultural development. Another possibility is absolute flatness. As Chayka explains, “flatness” refers to the homogenisation and reduction of culture, where “the least ambiguous, least disruptive, and perhaps least meaningful pieces of culture are promoted the most” (Chayka 2024, 17). The averageness of culture is calculated to match the “lowest common denominator” (ibid.). Since power is often profit-driven and seeks maximum engagement, the aesthetic that appeals to the widest audience prevails. The popular becomes more popular, the invisible becomes more invisible (ibid., 45), and existing power structures are reinforced. It becomes challenging to “separate the nature of something, or its reality, from its popularity” (ibid., 132). Popularity, amplified by algorithmic recommendation, “often gets confused for meaning or significance” (ibid.).

Unfortunately, art that appeals to a broad audience often lacks nuance and must remain "safe" and not overly experimental, relying on algorithmic evaluations of past successes. This trend is evident in the emphasis on the “cinematic” quality as a guiding principle for AI art. The desirable outputs described by Altman or Murati are not intended to challenge audiences or reframe the horizon of possibility in the minds of the users. Instead, they promote the numbing “beautiful,” “smooth,” and “cinematic” aspects of production. The popularity of unquestionably cinematic Marvel movies (part of the Marvel *Cinematic* Universe) demonstrates how easily consumable content, based on an unchallenging lens, can be.

Another “flattening” feature of AI creation is its iterative capacity. Our reality is already overflowing with content, and generative AI expands the quantity of cultural products exponentially. This provides a variety of choices. As I have shown, not only can we watch the Harry Potter saga, but we can also watch it in the style of Balenciaga, Italy, or the Vietnam War (@demonflyingfox 2023a; 2023b; 2024a). However, “the endless array of options presented by algorithmic feeds often instils a sense of meaninglessness” (Chayka 2024, 86). Unlimited

access to any piece of content makes it difficult to spend time with and appreciate any one piece. As Verbeek explains, “whenever there are a large number of identical items, no real significance can be attached to any particular one” (Verbeek 2005, 28). If all styles, characters, and techniques are reduced to mere technical variables for prompts in the creation of endlessly remixed culture, it is difficult to imagine that the resulting experience could be truly aesthetically or emotionally transformative. “Harry Potter by Balenciaga” garnered over 12 million views (@demonflyingfox 2023a). However, out of the dozens of similar creations that followed, few surpassed 2 million views (@demonflyingfox 2024b). The remixes quickly lost their viral potency. Of course, @demonflyingfox’s videos are not produced with Hollywood-level budgets or powerful marketing. However, by comparing this situation to the critique Disney faces for its reliance on live-action remakes rather than creating new stories (McCoy 2022), it can be speculated that audiences will not enjoy iterative or infinitely remixed cultural productions. This aversion may intensify as generative AI makes such productions even more easily producible for entertainment industry giants.

Possibly the biggest problem with algorithmically flattened culture goes beyond whether it is aesthetically or emotionally challenging. The problem is power. With the scale at which AI operates, it is worryingly easy for powerful agents to stay powerful. Access to widely applied, centralised models allows for “a self-perpetuating system that privileges the already-powerful” (Bones et al. 2021, 28). Creating something truly new with AI is already challenging. However, even if a creator manages to produce something novel, it could be suppressed if it threatens the popularity of existing cultural outputs that benefit powerful agents. This lack of access to virality means that, in the digital era, such creations could essentially cease to exist.

Human, the Meaning-Maker

Generative AI technologies are likely to transform the practices and perceptions of both creators and consumers (Kudina and Verbeek 2019, 295). Through the technologically-induced process of framing the eyes, these technologies will produce entirely new subjects. As the use of technology becomes normalised and users' perceptions shift towards (post-)algorithmic seeing, the understanding of what constitutes art, the role of humans in the creative process, and the technology itself is likely to change significantly.

For creators, the introduction of AI tools is expected to substantially alter traditional artistic workflows. The ease and efficiency provided by AI could democratise art creation, enabling even those with limited technical skills to produce high-quality work, which could lead to a more diverse and inclusive artistic community. Technological stabilities impact artistic choices and set limits on human imagination. While they can open up new experiences and possibilities, they also create new challenges and limitations. AI might expand aesthetic possibilities, allowing for innovative content that was previously unattainable, but it could also undermine the authenticity of human-made culture.

Novel functionalities, such as explicit prompting, will shift the importance from technical skills to rich cultural capital or uniquely human skills such as the use of humour or metaphors. The ability of users to incorporate new

technologies into their practices depends on the human-technology mutual constitution and perceptual frameworks. A culture guided by the notion of “cinematic” and limited by the data-based nature of generative technology will shape usage patterns, potentially leading to changes in perception and practices, and the displacement of human artists.

On the consumer side, generated content is likely to reshape how people engage with and consume media. AI-assisted personalisation promises media experiences tailored to individual preferences, making cultural consumption more engaging and satisfying. Consumers will encounter content that closely aligns with their tastes, potentially enriching their cultural experiences. However, such super-personalisation can pave the way for new types of advertising and even the fragmentation of shared reality. The ways of seeing assumed by users will be as diverse as their cultural exposure. At the same time, the technology's scale, guided by the numbing notions of “smooth” and “cinematic” and the development of iterative remix culture, can lead to homogenisation. Algorithms might prioritise mainstream trends and popular aesthetics, thereby reducing cultural diversity and exposure to novel or challenging works.

Technology and humans engage in a process of co-shaping. As powerful as an artefact can be, it is humans who are in the position of meaning-makers, deciding the social, moral, and aesthetic positions of the technology. Technological companies have a vested interest in influencing the sociotechnical imaginary that emerges. The dangers of AI are not in the technology itself but rather in the narratives surrounding it—especially those constructed by powerful agents. Profit-driven AI companies will try to control how the technology is perceived and used and what kinds of stabilities are established. The nature of AI sustains existing power structures, and it is likely that notions of human imagination expansion or discovery of new artistic avenues will not guide the culture constructed with AI. Instead, the guiding principle in capitalism is profit.

While AI integration may lead to the loss of certain human elements and a revolution in the job market, it also opens up new opportunities for enrichment and expression. The co-shaping interplay between humans and AI in the cultural domain and the emergent ways of seeing will require a reconsideration of existing practices and perceptions as creators and consumers navigate the reality co-shaped by algorithmic technologies. Susan Sontag's observation about the power of photographic seeing is pertinent here: “photography has succeeded in somewhat revising, for everybody, the definitions of what is beautiful and ugly” (Sontag 2008, 27). The definition or meaning of a phenomenon is revised by technology but not categorically established by it. The production of meaning is a human domain. It may be mediated by technology, but art, creativity, beauty, and other aesthetic or moral phenomena “are always relative to human evaluation” (Farina et al. 2024, 10). This meaning-making role is a key human faculty in the AI era.

Schmidt and Loidolt observed that machines can participate in making-meaning “for us” but they themselves are blind to that meaning (Schmidt and Loidolt 2023, 23). With technologies lacking the skill of meaning-making, that role remains firmly within human agency. Generative AI technologies are highly iterative, but it is the user who

judges the results as bad, good enough, or a “basis for further runs” (Magnus 2023, 4). Wellner agrees: “it is difficult for algorithms to decide which variation is meaningful. They simply produce more and more variations” (Wellner 2022, 201). She also asserts that “the logic of AI leaves the production of meaning to humans” (ibid.). It is the artist, designer, and prompter who decide “how to proceed at each step” and “the final image ultimately relies on their beliefs and attitudes” (Magnus 2023, 4). While technology might guide its use and technological companies might try to control its perception, it is important to keep in mind that it is within the human role to establish the meaning and stabilities of the technologies they use. In order to fulfil our meaning-making role we need to have a profound understanding of not only how the technology functions, but how it exists in a mutually constituting relationship to us.

5. Conclusion

The advent and rapid proliferation of visual generative technologies have initiated a profound transformation in human perceptions, habits, and experiences. This thesis has examined the multifaceted impact of these technologies through a postphenomenological lens, emphasising the co-constitutive relationship between humans and technology. A central insight from this research is the concept of "framing the eyes." Informed by the non-neutral transformative power of artefacts, it describes how gradual exposure to a specific technology transforms its users. In context of generative AI, this process involves the gradual habitualisation of users to AI-generated content, which transforms their perceptual frameworks and practices. Over time, as these technologies become deeply integrated into everyday life, what can be observed is the emergence of "(post-)algorithmic seeing"—a form of human perception that has been mediated by synthetic outputs of generative AI and lead to shifts in visual and aesthetic frameworks, as well as the ways in which these technologies shape our understanding of reality.

An important insight of this study was that with the emergence of new ways of seeing, some aspects of previous perceptual frameworks need to be discarded. In the context of algorithmic seeing, the notion of photographic believability, inherited from the photographic seeing conceptualised by Sontag, is no longer a reliable perceptual assumption. Generative AI, with its growingly unprecedented level of realism, challenges the perceptual frameworks of the pre-AI era.

The transformative power of artefacts plays a crucial role in understanding the impact of generative technologies on our perception. Technologies like Sora actively shape human experiences and societal norms, highlighting the bidirectional influence between humans and technology. As generative AI becomes more widely used, entirely new subjects will emerge from the co-shaping relation between humans and technologies.

From the perspective of a creator, multiple stabilities that come with introduction of the generative technology into visual culture can revolutionise artistic workflows. AI can democratise art creation and expand human creativity. The reliance on rich cultural capital and uniquely human skills, such as humour and metaphor, will become increasingly important as explicit prompting shifts the emphasis away from technical prowess. However, the data-based nature of the technology, reduction of cultural artefacts and artistic styles into prompt variables and the “cinematic” quality of outputs becoming a leading notion of generated outputs, might lead to a homogenised and unstimulating culture, driven not by artistic development, but profit-driven goals of AI and entertainment companies.

On the consumer side, AI-generated content is set to transform media engagement. AI-assisted personalization could make cultural consumption more engaging by tailoring content to individual preferences. However, this super-personalization carries risks, including the potential erosion of shared reality and the introduction of

subtle advertising that blurs the line between art and sales pitches. Additionally, the homogenization of culture through algorithmic prioritisation of mainstream trends threatens cultural diversity and exposure to novel or challenging works.

The challenges posed by AI are not inherent to the technology itself but stem from the narratives constructed around it, especially by powerful entities. Although initial user reactions to AI were marked by awe and fascination, a growing sense of disappointment and disillusionment has emerged regarding its capabilities. Technological companies continue to shape sociotechnical imaginaries, develop fabricated sublime, and engage in aesthetic quality dictation to establish influential stabilities that benefit the company. However, users do not always align with OpenAI's narratives. It is crucial to remain vigilant and critically engage with these technologies to ensure they serve humanity's best interests rather than merely those of the companies that produce them.

The shift in visual perception necessitates new forms of visual literacy and critical engagement. The subjects produced through this technologically-induced framing process reflect a dynamic interplay between human perception and praxis and technological mediation. The role of humans as meaning-makers remains crucial, yet in the context of late-stage capitalism, there is a risk that technology producers will unduly influence how these technologies are perceived and used.

We find ourselves at the beginning of a process with little insight into how these new-to-the-world technologies will shape the future (Kudina and Verbeek 2019, 296). In the years to come, the role of philosophers of technology lies in careful observation of the shifting relationship between AI technologies and human beings, especially the ways in which humans give meaning to those technologies (Verbeek 2016, 3). Future research should focus on developing frameworks for responsible AI use, fostering algorithmic literacy, and exploring the long-term impacts of these technologies on human perception and culture. The co-constitutive relationship between humans and technology, a central tenet of postphenomenology, is evident in the evolving dynamics of visual culture. As we become increasingly reliant on these technologies, it is imperative to consider how they mediate our interactions with the world and the kinds of subjects they produce. Only by critically engaging with these issues can we ensure that the evolution of visual generative technologies contributes to a more informed and equitable society.

References

@0xDesigner. 2023. "this is both magnificent and horrific" Twitter, March 28, 2023. https://x.com/0xDesigner/status/1640751951251144705

@adesigne. 2024. "The first SORA music video." *Reddit*, May 2, 2024. https://www.reddit.com/r/ChatGPT/comments/1cip21x/the_first_sora_music_video/.

@asdloller. 2023. Comment on"Harry Potter by Balenciaga." YouTube video. March 15, 2023. https://www.youtube.com/watch?v=iE39q-IKOzA

@atheneite. 2023. Comment on"Harry Potter by Balenciaga." YouTube video. March 15, 2023. https://www.youtube.com/watch?v=iE39q-IKOzA

@Beersnake_21. 2023. "I will never sleep again." Twitter, March 28, 2023. https://x.com/bearsnake_21/status/1640568000620417024

@casualblogger3810. 2023. Comment on"Harry Potter by Balenciaga." YouTube video. March 15, 2023. https://www.youtube.com/watch?v=iE39q-IKOzA

@chaindrop. 2023. "Will Smith eating spaghetti." *Reddit*, March 25, 2023. https://www.reddit.com/r/StableDiffusion/comments/1244h2c/comment/jdylwwq/?context=3.

@checkthreetimes. 2023. "This is quite disturbing" Twitter, March 29, 2023. https://x.com/checkthreetimes/status/1640901238945357824

@Crafty_Enthusiasm_99. 2023a. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@Crafty_Enthusiasm_99. 2023b. Comment on "Harry Potter by Balenciaga." *Reddit*, March 22, 2023. https://www.reddit.com/r/videos/comments/11ykbrg/harry_potter_by_balenciaga/.

@demonflyingfox. 2023a. "Harry Potter by Balenciaga." YouTube video. March 15, 2023. https://www.youtube.com/watch?v=iE39q-IKOzA

@demonflyingfox. 2023b. "Harry Potter but in Italy." YouTube video. June 3, 2023. https://www.youtube.com/watch?v=AN8FnohbJcw

@demonflyingfox. 2024a. "Harry Potter and the Vietnam War." YouTube video. June 6, 2024. https://www.youtube.com/watch?v=tEnrTVSbMro

@demonflyingfox. 2024b. *@demonflyingfox YouTube Channel*. YouTube. https://www.youtube.com/@demonflyingfox/videos.

@DPSonly. 2023. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@EddyScamArtist. 2024. “Probably gonna need to start figuring out UBI or some form of it as a society sooner rather than later lol.” Twitter, February 15, 2024. https://twitter.com/EddyScamArtist/status/1758229665335144896

@EddyVGG. 2023. "this is an atrocity pls reconsider your existence" Twitter, March 28, 2023. https://x.com/EddyVGG/status/1640809624868642816.

@ElectronicHouseFlash. 2023. Comment on"Harry Potter by Balenciaga." YouTube video. March 15, 2023. https://www.youtube.com/watch?v=iE39q-IKOzA

@ent4151. 2024. “What’s that magic?” Twitter, February 15, 2024. https://x.com/ent4151/status/1758200505933631790

@GaaraSama83. 2023. Comment on "Harry Potter by Balenciaga." *Reddit*, March 22, 2023. https://www.reddit.com/r/videos/comments/11ykbrg/harry_potter_by_balenciaga/.

@Kanute3333. 2024. "Comment on ’The first SORA music video.’" *Reddit*, May 2, 2024. https://www.reddit.com/r/ChatGPT/comments/1cip21x/the_first_sora_music_video/.

@luminous_lead. 2023. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@midnightmiragemusic. 2024. "Comment on ’The first SORA music video.’" *Reddit*, May 2, 2024. https://www.reddit.com/r/ChatGPT/comments/1cip21x/the_first_sora_music_video/.

@monsantobreath. 2023a. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@monsantobreath. 2023b. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@nathanstolen. 2023. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@nolaconnor. 2023. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@no-one-4845. 2024. "Comment on 'The first SORA music video.'" *Reddit*, May 2, 2024. https://www.reddit.com/r/ChatGPT/comments/1cip21x/the_first_sora_music_video/.

@Olney1Ben. 2024. "For the first time ever, we can literally recreate our own nightmares." Twitter, February 15, 2024. https://twitter.com/Olney1Ben/status/1758241813997981796.

@RationalEye. 2024. "Sam can't make money unless he keeps pitching fancy-sounding tech ideas to big-money donors He doesn't really care if, you know, those ideas a totally make sense... or if anyone even actually wants them" Twitter, April 5, 2024. https://x.com/RationalEye/status/1776124057408426304.

@supervegeta101. 2023. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@TheAvatarMovies. 2024. "Wow. This is going to change everything." Twitter, February 15, 2024. https://x.com/TheAvatarMovies/status/1758212028005552376.

@theseawillrise. 2023. Comment on"Harry Potter by Balenciaga." YouTube video. March 15, 2023. https://www.youtube.com/watch?v=iE39q-IKOzA

@trippy_art_special. 2023. "The Pope Drip." *Reddit*. March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

@TTEcclesBrown. 2023. "This is amazing!" Twitter, March 29, 2023. https://x.com/checkthreetimes/status/1640901238945357824

@UnequalBull. 2024. "Comment on 'The first SORA music video.'" *Reddit*, May 2, 2024. https://www.reddit.com/r/ChatGPT/comments/1cip21x/the_first_sora_music_video/.

@user-lm6yk4hh5z. 2023. Comment on"Harry Potter by Balenciaga." YouTube video. March 15, 2023. https://www.youtube.com/watch?v=iE39q-IKOzA

Altman, Sam (@sama). 2024a. "here is sora, our video generation model:http://openai.com/sora today we are starting red-teaming and offering access to a limited number of creators. @_tim_brooks @billpeeb @model_mechanic are really incredible; amazing work by them and the team. remarkable moment." Twitter, February 15, 2024. https://x.com/sama/status/1758193609927721350.

Altman, Sam (@sama). 2024b. "movies are going to become video games and video games are going to become something unimaginably better." Twitter, April 5, 2024. https://x.com/sama/status/1776083954786836979?lang=en

Anthis, Jacy Reese (@jacyanthis). 2023. "This looks silly and absurd, but it's actually what text-to-image looked like 2 years before the, 'Wow, this is amazing!' events of 2021. Maybe we will see amazing text-to-video by 2025." Twitter, March 29, 2023. https://x.com/jacyanthis/status/1641087957175083008

ArmchairReditor. 2023. Comment on "The Pope Drip." *Reddit*, March 26, 2023. https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/.

Benjamin, Walter. 2008. The Work of Art in the Age of Mechanical Reproduction. Translated by J.A. Underwood. London: Penguin Books.

Berger, John. 1972. *Ways of Seeing*. London: Penguin Books.

Blok, Vincent. 2022. "The Role of Human Creativity in Human-Technology Relations." Philosophy & Technology 35 (3): 59. https://doi.org/10.1007/s13347-022-00559-7.

Boden, Margaret A. 2004. *The Creative Mind: Myths and Mechanisms*. London: Routledge.

Bones, Helen, Susan Ford, Rachel Hendery, Kate Richards, and Teresa Swist. 2021. "In the Frame: The Language of AI." *Philosophy & Technology* 34 (S1): 23–44. https://doi.org/10.1007/s13347-020-00422-7.

Bridle, James. 2019. "New Ways of Seeing: Can John Berger's Classic Decode Our Baffling Digital Age?" *The Guardian*, April 16, 2019. https://www.theguardian.com/books/2019/apr/16/new-ways-of-seeing-can-john-bergers-classic-decode-our-baffling-digital-age.

Broderick, Ryan (@broderick). 2023. "I think Balenciaga pope might be the first real mass-level AI misinformation case. Kinda cool to watch tbh." Twitter, March 26, 2023. https://twitter.com/broderick/status/1640039713649094656.

Brownlee, Marques. 2024. “AI Generated Videos Just Changed Forever.” Youtube video, 12:01. February 16, 2024. https://www.youtube.com/watch?v=NXpdyAWLDas

Bucher, Taina. 2017. “The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms.” Information, Communication & Society 20 (1): 30–44. https://doi.org/10.1080/1369118X.2016.1154086.

Burnham, Douglas. n.d. “Immanuel Kant: Aesthetics.” Internet Encyclopedia of Philosophy. https://iep.utm.edu/kantaest/#SH2c

Celis Bueno, Claudio, Pei-Sze Chow and Ada Popowicz. 2024. “Not ‘what’, but ‘where is creativity?’: towards a relational-materialist approach to generative AI.” *AI & Society (2024)*. https://doi.org/10.1007/s00146-024-01921-3.

Chayka, Kyle. 2022. “Have iPhone Cameras Become Too Smart?” *The New Yorker*, March 18, 2022. https://www.newyorker.com/culture/infinite-scroll/have-iphone-cameras-become-too-smart

Chayka, Kyle. 2024. *Flatland: How Algorithms Flattened Culture*. New York: Bloomsbury Publishing.

Chesher, Chris, and César Albarrán-Torres. 2023. “The Emergence of Autolography: The ‘Magical’ Invocation of Images from Text through AI.” *Media International Australia* 189 (1): 57–73. https://doi.org/10.1177/1329878X231193252.

Chhetri, Vivek. 2024. “OpenAI’s Sora on the rise: Inside Sam Altman’s $7 Trillion AI investment quest.” *Tech Funding News*, February 27, 2024. https://techfundingnews.com/openais-sora-on-the-rise-inside-sam-altmans-7-trillion-ai-investment-quest/

Coeckelbergh, Mark. 2017. “Language and Technology: Maps, Bridges, and Pathways.” AI & Society 32: 175–189. https://doi.org/10.1007/s00146-015-0604-9.

Coeckelbergh, Mark. 2023. “The Work of Art in the Age of AI Image Generation: Aesthetics and Human-Technology Relations as Process and Performance.” Journal of Human-Technology Relations 1. https://doi.org/10.59490/jhtr.2023.1.7025.

Delahaye, Jean-Paul. 2020. "For Math Fans: a Hitchhiker's Guide to the Number 42." *Scientific American*, September 21, 2020. https://www.scientificamerican.com/article/for-math-fans-a-hitchhikers-guide-to-the-number-42/.

DePew, Charlie (@CharlieDePew). 2024. "Film industry (in its current form) is so fucked." Twitter, February 16, 2024. https://x.com/CharlieDePew/status/1758273325489656279.

Deshpande, Sanika (@sanika_desh). 2024. "This is going to disrupt the current VFX industry for sure. Producers will hug you :)." Twitter, February 23, 2024. https://x.com/sanika_desh/status/1760908958754353561

Dhesi, Aman (@amansplaining). 2024. "The future is starting to look like the future." Twitter, February 15, 2024. https://twitter.com/amansplaining/status/1758206778955751544.

Di Placido, Dani. 2023. "Why Did 'Balenciaga Pope' Go Viral?" *Forbes*, March 27, 2023. https://www.forbes.com/sites/danidiplacido/2023/03/27/why-did-balenciaga-pope-go-viral/.

Di Placido, Dani. 2023. "'Demonic' AI-Generated 'Will Smith Eating Spaghetti' Clip Goes Viral." *Forbes*, April 3, 2023. https://www.forbes.com/sites/danidiplacido/2023/04/03/will-smith-eating-spaghetti-is-the-latest-ai-generated-eldritch-abomination/.

Eisner, Lotte H. 1983. *The Haunted Screen: Expressionism in the German Cinema and the Influence of Max Reinhardt*. Berkeley: University of California Press.

Elias, Michelle and Razik, Navee. 2023. "'Balenciaga Pope' might not have been real. But its impact is." *SBS News*, March 29, 2023. Accessed May 22, 2024. https://www.sbs.com.au/news/the-feed/article/balenciaga-pope-might-not-have-been-real-but-its-impact-is/61v1i9h3x.

Farina, Mirko, Andrea Lavazza, Giuseppe Sartori, and Witold Pedrycz. 2024. "Machine Learning in Human Creativity: Status and Perspectives." AI & Society (2024). https://doi.org/10.1007/s00146-023-01836-5.

Ford-Monroe, Toby (@towbew). 2024. "This is unreal." Twitter, February 15, 2024. https://x.com/towbew/status/1758207087744516406

Fryant, Christopher (@cfryant). 2024. "Nah, movies are 'content' now." Twitter, April 5, 2024. https://twitter.com/cfryant/status/1776088849195909321

Gerlek, Selin, and Sebastian Weydner-Volkmann. 2022. "Self-Tracking and Habitualization. (Post)-Phenomenological and Pragmatist Perspectives on Reflecting Habits with the Help of Digital Technologies." In: Selin Gerlek, Sarah Kissler, Thorben Mämecke, and Dennis Möbus (eds.): Von Menschen und Maschinen: Mensch-Maschine-Interaktionen in digitalen Kulturen, 136–149. https://doi.org/10.57813/20220623-152405-0.

Gregor, Ulrich and Enno Patalas. 1962. *Geschichte des Films (History of Film)*. Reinbek bei Hamburg: Rowohlt.

Grundhauser, Eric. 2016. "Did a Silent Film About a Train Really Cause Audiences to Stampede?" *Atlas Obscura*, November 3, 2016. https://www.atlasobscura.com/articles/did-a-silent-film-about-a-train-really-cause-audiences-to-stampede.

Halford, Justin (@Justin_Halford_). 2024. "Holy shit. All details addressed. Long term coherence.Bearish Pixar and Disney. Bullish the owners of compute." Twitter, February 15, 2024. https://x.com/Justin_Halford_/status/1758201153416724929.

Hard Fork. 2024. "Which face is A.I.? | Clip. " YouTube video, 4:08. January 26, 2024. https://www.youtube.com/watch?v=netLOj3g4LM.

Hassid, Ruben (@RubenHssd). 2024. "AI goes too fast. Let us absorb for a sec." Twitter, February 15, 2024. https://x.com/RubenHssd/status/1758201388260008229

Heaven, Will Douglas. 2024. "OpenAI teases an amazing new generative video model called Sora." *MIT Technology Review*, February 15, 2024. https://www.technologyreview.com/2024/02/15/1088401/openai-amazing-new-generative-ai-video-model-sora/.

Heikkilä, Melissa. 2022. "This artist is dominating AI-generated art. And he's not happy about it." *MIT Technology Review*, September 16, 2022. https://www.technologyreview.com/2022/09/16/1059598/this-artist-is-dominating-ai-generated-art-and-hes-not-happy-about-it/

House, Patrick. 2024. "The Lifelike Illusions of AI." *The New Yorker*, March 19, 2024. https://www.newyorker.com/science/annals-of-artificial-intelligence/the-lifelike-illusions-of-ai.

Huang, Kalley. 2023. "Why Pope Francis Is the Star of A.I.-Generated Photos." *The New York Times*, April 8, 2023. https://www.nytimes.com/2023/04/08/technology/ai-photos-pope-francis.html.

Ihde, Don. 1990. *Technology and the Lifeworld: From Garden to Earth.* Bloomington: Indiana University Press.

Ihde, Don. 1999. "Expanding Hermeneutics." In *Hermeneutics and Science*, edited by Márta Fehér, Olga Kiss, and László Ropolyi, 206:345–51. Boston Studies in the Philosophy of Science. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-015-9293-2_28.

Ihde, Don. 2009. *Postphenomenology and Technoscience: The Peking University Lectures*. New York: State University of New York Press.

Iyengar, Aishwariya, Rutvij Kulkarni and T.N.C Vidya. 2015. "Koinophilia and human facial attractiveness". *Resonance*. 20, 311–319 (2015). https://doi.org/10.1007/s12045-015-0187-2.

Jarow, Oshan. 2023. "How fake AI images can expand your mind." *Vox*, March 30, 2023. https://www.vox.com/future-perfect/23661673/pope-puffer-coat-generative-ai-midjourney-imagination.

Jasanoff, Sheila, and Kim, Sang-Hyun . 2009. "Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea." Minerva 47 (2): 119–46. https://doi.org/10.1007/s11024-009-9124-4.

Jennings, Rebecca. 2023. "AI art freaks me out. So I tried to make some." *Vox*, March 20, 2023. https://www.vox.com/culture/23678708/ai-art-balenciaga-harry-potter-midjourney-eleven-labs.

John, Daniel. 2024."Mind-blowing videos reveal how AI has improved in just 1 year." *Creative Bloq*, February 20, 2024. https://www.creativebloq.com/news/ai-video-progress

Karasek, Hellmuth. 1994. "Lokomotive Der Gefühle (Locomotive of Emotions)" *Der Spiegel*, December 26, 1994. https://www.spiegel.de/kultur/lokomotive-der-gefuehle.

KettleBellDan. 2024. "Can't movies just stay movies. I like those." Twitter, April 5, 2024. https://twitter.com/KettlebellDan/status/1776088575962141134

Kudina, Olya, and Peter-Paul Verbeek. 2019. “Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy.” Science, Technology, & Human Values 44 (2): 291–314. https://doi.org/10.1177/0162243918793711.

Lee, Nicol Turner and Natasha White. 2024. “How OpenAI’s Sora hurts the creative industries.” *Brookings*, May 15, 2024. https://www.brookings.edu/articles/how-openais-sora-hurts-the-creative-industries/

Levy, Steven. 2024. "OpenAI’s Sora Turns AI Prompts Into Photorealistic Videos" *Wired*, February 15, 2024. https://www.wired.com/story/openai-sora-generative-ai-video/.

Loiperdinger, Martin. 2004. "Lumière's Arrival of the Train: Cinema's Founding Myth." *The Moving Image: The Journal of the Association of Moving Image Archivists 4, no. 1 (2004): 89–118. http://www.jstor.org/stable/41167150.*

Lumière, Auguste, and Louis Lumière. 1896. *Arrivée d'un train à la Ciotat (Arrival of a Train at La Ciotat).* YouTube video. Posted by Nubsy, October 31, 2014. https://www.youtube.com/watch?v=MT-70ni4Ddo&themeRefresh=1.

Magnus, P. D. 2023. “Generative AI and Photographic Transparency.” AI & Society (2023). https://doi.org/10.1007/s00146-023-01817-8.

Manovich, Lev. 2023. “AI image and Generative Media: Notes on Ongoing Revolution.” Manovich Blog. http://manovich.net/content/04-projects/167-artificial-aesthetics-book/manovich-ai-image-and-generative-media.pdf

McCoy, Joshua Kristian. 2022. “Why Do People Hate Disney's Live-Action Remakes?” GameRant, August 8, 2022. https://gamerant.com/disney-live-action-remakes-hate/.

Metz, Cade. 2024. "OpenAI Unveils A.I. That Instantly Generates Eye-Popping Videos." *The New York Times*, February 15, 2024. https://www.nytimes.com/2024/02/15/technology/openai-sora-videos.html.

Metz, Cade and Tripp Mickle. 2024. “OpenAI Completes Deal That Values the Company at $80 Billion.” *The New York Times*, February 16, 2024. https://www.nytimes.com/2024/02/16/technology/openai-artificial-intelligence-deal-valuation.html

Midjourney. 2024. Midjourney Discord Server. Accessed June 11th, 2024. https://discord.gg/midjourney

Morley, Simon. 2021. "A Short History of the Sublime." *MIT Press Reader*, March 22, 2021. https://thereader.mitpress.mit.edu/a-short-history-of-the-sublime/.

Morrison, Ryan. 2024. “I just tried MidJourney's new consistent character feature — here's how it looks.” *Tom’s Guide*, March 12, 2024. https://www.tomsguide.com/ai/ai-image-video/i-tried-midjourneys-new-consistent-character-feature-heres-how-it-turned-out

Murati, Mira. 2024. Interview by Joanna Stern. *The Wall Street Journal*. YouTube video, 10:38. March 13, 2024. https://www.youtube.com/watch?v=mAUpxN-EIgU.

Museum of Modern Art (MoMA). n.d. "*The Arrival of a Train (L'arrivée d'un train en gare de La Ciotat).*" Accessed May 22, 2024. https://www.moma.org/collection/works/307080.

Newton, Casey and Kevin Roose. 2024. “Is A.I. Already Taking Jobs? +A Filmmaker Tries Sora + The XZ Backdoor Caper.” Produced by Rachel Cohn and Whitney Jones. *Hard Fork*. April 5, 2024. Podcast, website, 1:04:24. https://www.nytimes.com/2024/04/05/podcasts/hardfork-ai-jobs-sora-xz.html?

Nyce, Caroline Mimbs. 2023. “AI Has a Hotness Problem.” *The Atlantic*, October 24, 2023. https://www.theatlantic.com/technology/archive/2023/10/ai-image-generation-hot-people/675750/

OpenAI. 2024a. "Sora: Advanced AI Video Model." OpenAI Blog, February 15, 2024. Accessed May 22, 2024.https://openai.com/index/sora/.

OpenAI. 2024b. "Careers at OpenAI." OpenAI Blog. Accessed May 22, 2024. https://openai.com/careers/.

OpenAI. 2024c. "Sora: First Impressions." OpenAI Blog, March 25, 2024. Accessed May 22, 2024. https://openai.com/index/sora-first-impressions/

OpenAI. 2024d. “air head · Made by shy kids with Sora.” Youtube video, 1:21. April 5, 2024. https://www.youtube.com/watch?v=9oryIMNVtto

Pequeño, Antonio IV. 2024. "OpenAI Reveals ‘Sora’: AI Video Model Capable Of Realistic Text-To-Video Prompts." *Forbes*, February 15, 2024. https://www.forbes.com/sites/antoniopequenoiv/2024/02/15/openai-reveals-sora-ai-video-model-capable-of-realistic-text-to-video-prompts/.

Perrigo, Billy. 2023. "How to Spot an AI-Generated Image Like the 'Balenciaga Pope'." *Time*, March 28, 2023. Accessed May 22, 2024. https://time.com/6266606/how-to-spot-deepfake-pope/.

Pritchard, Stephern. 2023. "Artwashing: Social Capital & Anti-Gentrification Activism." *Colouring in Culture Blog*, June 13, 2017. https://www.colouringinculture.org/blog/uncategorized/artwashingsocialcapitalantigentrification.

PromptJungle. 2023. "Harry Potter By Balenciaga | Step by Step Tutorial" YouTube video. April 3, 2023. https://www.youtube.com/watch?v=TGD8zKvRxc4&t=207s.

Rosenberger, Robert, and Peter-Paul Verbeek. 2015. *Postphenomenological Investigations: Essays on Human–Technology Relations*. Lanham: Lexington Books.

Runquist, Karsten. 2024. "AI and Filmmaking." YouTube video, 10:24. April 16, 2024. https://youtu.be/PROQJcKd678?si=I1ocA1tvd8K-g2fq.

Sanders, Nathan E. and Schneier, Bruce. 2024. "Let's Not Make the Same Mistakes with AI That We Made with Social Media." *MIT Technology Review*, March 13, 2024. https://www.technologyreview.com/2024/03/13/1089729/lets-not-make-the-same-mistakes-with-ai-that-we-made-with-social-media/?truid=&mc_cid=e52ab14411.

Schmidt, Philipp and Loidolt, Sophie. "Interacting with Machines: Can an Artificially Intelligent Agent Be a Partner?" *Philosophy & Technology.* 36, 55 (2023). https://doi.org/10.1007/s13347-023-00656-1.

Shreya, Manjari. 2023. "What is Banner Blindness? And the 7 Clever Ways To Avoid It." *Mastroke*, March 9, 2023. https://www.mastroke.com/blog/digital-marketing/what-is-banner-blindness-and-how-to-avoid-it-7-ways/

Sontag, Susan. 2008. *On Photography*. London: Penguin Books.

Strzemiński, Władysław. 2016. *Teoria widzenia (Theory of Seeing)*. Łódź: Muzeum Sztuki w Łodzi.

Tangermann, Victor. 2024. "Facebook Is Making Insane AI-Generated Fever Dreams Go Viral for Gullible Boomers." *Futurism*, March 19, 2024. https://futurism.com/facebook-ai-generated-images-viral

Trillo, Paul (@paultrillo). 2024. "First official commissioned music video made with @OpenAI Sora for @realwashedout This was an idea I had almost 10 years ago and then abandoned. Finally was able to

bring it to life. Watch the full video here https://vimeo.com/941713443." Twitter, May 2, 2024. https://twitter.com/paultrillo/status/1786044776745505209.

Verbeek, Peter-Paul. 2005. *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Philadelphia: Pennsylvania State University Press.

Verbeek, Peter-Paul. 2008. "Obstetric Ultrasound and the Technological Mediation of Morality: A Postphenomenological Analysis." Human Studies 31 (1): 11–26. https://doi.org/10.1007/s10746-007-9079-0.

Verbeek, Peter-Paul. 2015. "Beyond Interaction: A Short Introduction to Mediation Theory." Interactions 22 (3): 26–31. https://doi.org/10.1145/2751314.

Verbeek, Peter-Paul. 2016. "Toward a Theory of Technological Mediation:A Program for Postphenomenological Research." In: J.K. Berg O. Friis and Robert C. Crease, *Technoscience and Postphenomenology: The Manhattan Papers*. London: Lexington Books.

Vincent, James. 2023. "The swagged-out pope is an AI fake — and an early glimpse of a new reality." *The Verge*, March 27, 2023. https://www.theverge.com/2023/3/27/23657927/ai-pope-image-fake-midjourney-computer-generated-aesthetic.

Warzel Charlie. 2022. "What's really behind those AI art images?" *The Atlantic*, September 7, 2022. https://newsletters.theatlantic.com/galaxy-brain/6317de90bcbd490021b246bf/ai-art-dalle-midjourney-stablediffusion/\

Washed Out. 2024. "Washed Out - The Hardest Part (Official Video)" YouTube video, 4:02. May 2, 2024. https://www.youtube.com/watch?v=-Nb-M1GAOX8.

Wellner, Galit. 2014. "Multi-Attention and the Horcrux Logic: Justifications for Talking on the Phone While Driving." Techné: Research in Philosophy and Technology 18(1/2): 48–73. https://doi.org/10.5840/techne201432712

Wellner, Galit. 2022. "Digital Imagination: Ihde's and Stiegler's Concepts of Imagination." Foundations of Science 27 (1): 189–204. https://doi.org/10.1007/s10699-020-09737-2.

Wittingslow, Ryan. 2020. "Effing the Ineffable: The Sublime in Postphenomenology." Techné: Research in Philosophy and Technology 24 (3): 282–305. https://doi.org/10.5840/techne202082127.