Emanuele Arielli

Chapter 8. From Tools to Authors

Chapter 8 of *Artificial Aesthetics: Generative AI, Art and Visual Media* by Lev Manovich and Emanuele Arielli. Published October 27, 2024. 9,786 words. All book chapters: http://manovich.net/index.php/projects/artificial-aesthetics-book

> "If you heard someone playing the piano, would you ask: 'Is the piano the artist?' No. So, same thing here. Just because it is a complicated mechanism, it doesn't change the roles [...] The typewriter enables someone to write a book. For me, the keyboard enables me to write code, [...] there are neural networks involved that maybe you could say that they are my brushes that I learn to use."
> Mario Klingemann[1]

In philosophy of mind, we talk of "extended" mind referring to the fact that tools and technological advancement, including writing and memory devices, allowed to externalize and extend human mental processes in the outside environment. Technology enhances our "bounded rationality" (H. Simon) and extends our senses (M. McLuhan) that are limited by biological constraints. We have for instance limited memory, therefore writing and documents helped us to externalize and extend our capacity to recall. We have biologically limited visual acuity, but microscopes and telescopes allowed to amplify what could be visible to us. Similarly, our cognitive abilities to process numbers and data have upper limits, but calculators and computers have enabled us to surpass these boundaries.

Under this premise, machine learning and AI should be considered a further step in our *tool-making* ability in expanding humans' skills through devices. Therefore, their impact should be measured in their influence and contribution to human potentialities, not necessarily in their potential autonomy from human choice. The extended mind paradigm could be here coupled with an understanding of

technology not as a separate entity from human nature, but as a process of integration and augmentation between mind and technology.

Extended Aesthetics

It would be naïve to think that the human brain is suboptimal only in terms of memory and calculation. Other human faculties, such as imagination, perceptual sensitivity, emotional recognition and expression, and creativity, also have natural limits. Since these faculties are crucial in aesthetics, one could argue that humans possess not only "bounded rationality" but also "bounded aesthetic capacities." Although we can train and expand our aesthetic skills—both in appreciating and creating aesthetic objects— our biological limits mean we eventually reach a plateau, a "peak aesthetic sensitivity" and "peak creativity." Just as bounded rationality can be extended through external tools, our bounded aesthetic capabilities might also be enhanced and supplemented by tools that assist in the creative process. Our abilities to articulate ideas in writing may suffer from cognitive limitations as well: in this context, Large Language Models often come to our aid, suggesting responses to letters, helping to draft emails, and contributing to the clarity and persuasiveness of texts of all kinds. AI language models increasingly resemble a hidden assistant, providing clever responses and clear formulations, much like how Cyrano de Bergerac fed lines to Christian in the famous fictionalization of his life. These systems do the heavy lifting behind the scenes, granting users access to a vast array of knowledge and eloquence they might not possess on their own, thereby expanding human expressive possibilities.

Aesthetic has always been *extended*. All tools we use, from chisel for woodcarving or sculpting, to brushes for paintings to musical instruments and camera for photography, can be seen as extensions of our ability to create aesthetic artifacts. A brush can distribute chromatic pigments on a canvas in ways that we wouldn't be able to achieve with our hands alone. The simple act of sketching by hand on paper is a fundamental method of externalizing the images that arise in our minds, where both the productive and receptive sides work together in a continuous feedback cycle: since our working memory has limited capacity to retain an idea, we lay it down on paper. Our eyes then observe the sketch, allowing us to rework and develop the idea in an iterative and productive cycle between the eye, mind, sketching hand, and image on paper. Throughout history, the ability to create images has been one of the primary methods for externalizing memory and imagination, preserving them in tangible, enduring forms.

The tools that modern artists and designers use, like image editing software, computer-aided design (CAD) programs, and music production software, can be seen as modern extensions of human creativity as well. These technologies essentially distribute part of the aesthetic decision-making process outside of the artist's mind. A composer or writer stuck in their creative process might use more and more advanced systems that offer suggestions, evaluate alternative directions, test whether their ideas might receive a positive response from the public, and so on. One notable example is the case of composer David Cope in the early 1980s. He found himself struggling with severe composer's block while working on a commissioned opera. Rather than pushing through, he diverted his attention to developing a music composition program, a project that eventually evolved into what is now known as Experiments in Musical Intelligence, or EMI (often referred to as Emmy). EMI analyzes existing musical works in its database and generates new compositions in the same style, without simply copying the originals. Through this program, Cope has created thousands of works in diverse styles, including 5,000 Bach-likes chorales.

"Extended aesthetics" refers not only to the fact that tools broaden our creative possibilities, but also to the idea that sensibility, taste, intuition, and imaginative processes can be externalized as well. In this context, the extension through external devices affects not only production but also reception. Just as photographic reproduction has simultaneously extended and modified our perception of reality by creating new possibilities for artistic production, so too do the analysis and generation capabilities of new AI applications extend and modify our perception, as discussed in Chapter 6. Advanced systems of cultural analysis can deepen our understanding and aesthetic sensibility, for instance, by finding subtle associations or similarities between objects, comparing variants of similar artifacts, and detecting relevant details that we were previously unaware of. Various forms of artificial image post-production, translation, stylistic transfer, and morphing encourage us to view things through the lens of these transformative possibilities. By presenting new possibilities, these processes can change and refine our perception and taste, as is already evident with the content we encounter through recommendation systems.[2]
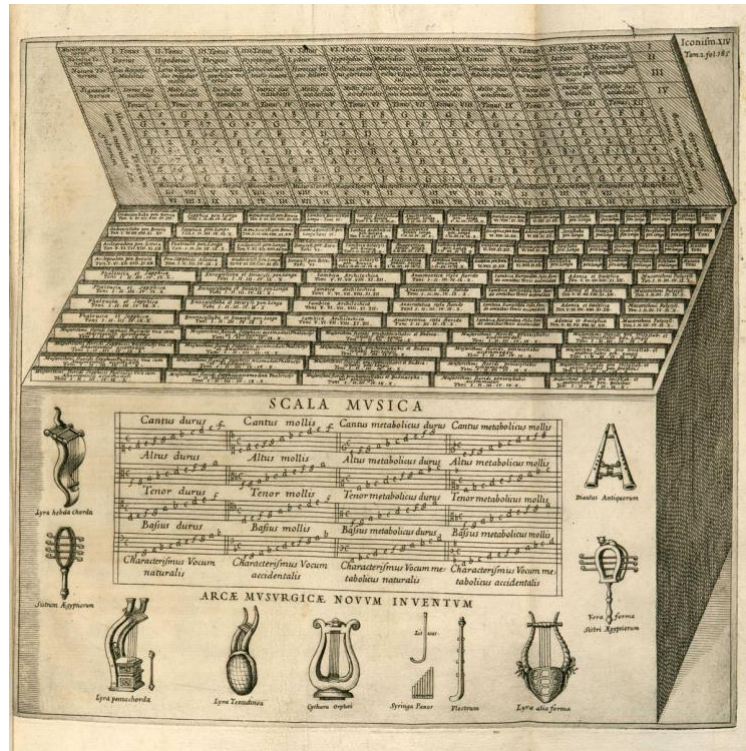
Chapter_08_Figure_01. *Arca Musarithmica*, Athanasius Kircher (1602-1680),

An early example of generative aesthetic device is the *Arca Musarithmica*, conceived by the German Jesuit Athanasius Kircher (1602-1680), a system based on tables and strips engraved with various musical elements, such as rhythms, melodic fragments, and harmonic progressions. By manually selecting and combining these elements and following a set of instructions provided by Kircher, it was possible to create compositions by adjusting mood, meter, and desired style. For many observers, this was the first algorithmic system for creative music generation. Since the user can randomly decide the parameters, the *Arca Musarithmica* also anticipates modern aleatory compositional techniques: the user, while adjusting general parameters, could not predict what kind of composition the system would ultimately generate.

By systematizing the creative process, Kircher's goal was to allow even non-musicians to compose music, as expressed in his 1650 treatise *Musurgia universalis*. Extending human capabilities allows individuals without specific skills to practice at a level that would otherwise only be possible for those with expertise. A generative device, in some ways, encapsulates expert skills codified in a tool external to our mind. In this sense, generative systems allow those who don't know how to draw to produce drawings according to their instructions, or those who don't know a language to produce a text in

that language. For some, this might represent a democratizing step, similar to how Walter Benjamin viewed the democratization of art through the possibilities of access created by the technical reproducibility of artworks.

As we enhance human faculties with devices, we still remain interested in assessing an individual's capabilities without such technological aids. For example, in a chess tournament between people, we are obviously interested in understanding a player's ability without the use of a chess program. Similarly, we measure the athletic qualities of a runner by making them run with their legs, not with a means of locomotion. While a motorcycle allows anyone to move quickly, only a few individuals achieve Olympic-level performances using just their own physical capabilities.

The concern here is that relying on technological extensions may lead to the atrophy of skills previously developed without them. For instance, the widespread use of pocket calculators has corresponded with a decrease in average mental and manual calculation abilities. Recent studies on the use of language models in schools show mixed results. While groups using these systems perform better in text composition, individual performance suffers when the technological aid is removed. Students using AI language models as a "crutch" may find their skill development impaired. [3]

As generative AI develops in fields like text, image, and music production, it raises the question of whether it's important for individuals to know how to excel without these tools. In fact, using them in art is often viewed as "cheating." This perspective is relatively new in creative contexts, except for certain professional photography competitions where digital postproduction is not allowed. Should we imagine poetry competitions where poets are not allowed to use language models? Or architecture awards where the use of generative systems is banned? Outside of specific competitive contexts, the field of cultural, aesthetic, and artistic production generally doesn't impose such creative constraints. If a musician's goal is to create the catchiest song possible, they may freely use systems that quickly generate musical variations. Similarly, in film or series production, algorithmic evaluation systems are tools used to reach maximum audience appeal. A platform like Netflix, for instance, has transformed movie and TV production with its data-driven approach. [4] By analyzing viewer preferences, they can predict which new shows will succeed and decide whether to continue existing series. This method examines specific elements that resonate with audiences, like certain scenes or episodes. The step from algorithmic data analysis to AI systems generating screenplay ideas is relatively short. As AI technology advances, future content creation

could plausibly involve AI-generated storylines and concepts based on the vast amounts of viewer preference data collected by streaming platforms.

From Tools to Agents

Advancements in AI are likely to evolve beyond the simple user-tool interaction. We're moving towards a more collaborative, dialogic, and iterative relationship between humans and machines. In this new paradigm, the machine's responses may not be easily predictable by the user. Instead, they will generate new ideas and directions, stimulating the user's own creativity and thought processes. One aspect of this technical evolution is the increasing decisional autonomy of devices. However, we often use the verb "decide" metaphorically when discussing machine actions. For instance, is a thermostat "deciding" when to stop heating based on a target temperature comparable to a self-driving car "deciding" to adjust its speed according to traffic conditions? There is no clear boundary between tools and entities with autonomous agency, just as there is no sharp distinction between a metaphorical and literal use of a term like "decide" (see Chapter 3). An attempt to classify the levels of autonomy in AI systems is, for example, the one suggested by Google DeepMind researchers, who propose a taxonomy of six levels.[5] At the lowest level (Autonomy Level 0: No AI), we find mere tools, such as a hammer or scissors, or in the artistic context, these could include analog tools like pencils and brushes or basic digital image editing software. At this level, humans do everything.

Moving up to Autonomy Level 1 (AI as a Tool), we encounter systems that automate secondary tasks while remaining under full human control, such as a thermostat regulating room temperature, software correcting text grammar, or performing translations. In the context of art, this could include machine learning-based tools for image enhancement, color correction, or style transfer. These tools augment human creativity but do not independently generate original content. For example, every modern phone is equipped with a system that helps a photographer by automatically adjusting exposure and color balance, but the artistic vision and composition remain firmly in the hands of the human artist. Even early generative programs based on well-defined computational procedures can be considered as belonging to this level. A. Michael Noll, a 1960s pioneer in computer art, programmed digital computers to mimic works by artists like Mondrian and Riley. His innovative approach combined specific algorithmic instructions with pseudo-random permutations.

At Autonomy Level 2 (AI as Advisor), AI takes on a substantive role, but only when invoked by a human. In the field of generative art, this could manifest as systems that can assist the artists guiding them through their process. For example, various platforms offer tools for music composition that leverage AI to enhance creativity. These tools can assist composers in exploring new musical ideas and structures, making the creative process more efficient and innovative. Machine learning systems that classify human emotions during music listening (applied, for example, by platforms like Spotify) can be used as an assistive tool for musicians to optimize their compositions.[6]

Progressing to Autonomy Level 3 (AI as Collaborator), the relationship between human and AI becomes more balanced, with interactive coordination of goals and tasks. In this case, the AI is capable of doing things that the human is not equipped for: think of a chess program guiding a human player's game. In the context of artistic creation, this level could be represented by systems that can engage in a creative process of mutual exchange with the human artist. The AI might propose compositions, color schemes, or stylistic choices, while the human provides feedback, refinements, and overall artistic direction. For example, a system like DALL-E or Midjourney generates images according to the artist's linguistic prompts, which the artist then refines, combines, or modifies according to their needs. The AI acts as a creative partner, offering suggestions and possibilities, but the human artist maintains primary control over the final artwork.

Autonomy Level 4 (AI as Expert) pushes the balance further towards AI dominance, with the AI system guiding the interaction while the human provides orientation, feedback, or performs secondary tasks. In aesthetics, this could manifest as highly sophisticated AI systems capable of generating entire artworks based on high-level concepts or themes provided by human curators. The human's role becomes more akin to that of an art director or curator, shaping the overall vision while the AI handles most of the creative execution.

These autonomy levels aren't rigid or mutually exclusive. AI systems in art often display traits from multiple levels, varying with the application and context. As AI evolves, these boundaries may blur further. Systems could shift between autonomy modes depending on the task and, crucially, how much control the human artist chooses to retain. Finally, at Autonomy Level 5 (AI as Agent), we encounter fully autonomous AI systems, a purely speculative prospect at present. This represents a theoretical endpoint where AI systems could conceive, create, and potentially even critique their own artworks. Importantly, "agency" and autonomy in this context don't refer to problematic concepts

like "free will" or "consciousness." Instead, they denote the capacity for autonomous goal-setting, goal execution, and the collective recognition of such autonomy. [7]

Art history provides several examples of the relationship between an artist and a "collaborator." In Renaissance workshops, masters worked alongside collaborators and apprentices, guiding the production of works and often contributing to key elements such as drawing, composition, and final touches. While apprentices and assistants frequently executed parts of these works, the finished pieces were usually attributed to the master, considered the creative force behind the work. An apprentice, typically a child, would start as an assistant with a purely "tool-like" function (mixing colors, preparing canvases), then gradually climb the autonomy hierarchy, progressing to collaborator and eventually becoming an autonomous master.

From the Renaissance through the 17th century and beyond, not all paintings from established workshops were entirely the master's handiwork. Artists like Giotto or Raphael gave ample space to workshop collaborators, while others, like Michelangelo, were reluctant to entrust work to others. Regardless of the assistants' autonomy levels, the master maintained overall creative control and authorship. The assistant was neither an expert nor an autonomous agent until leaving the master's workshop.

A similar dynamic exists in contemporary production of ready-mades and conceptual artworks. Here, the artist typically determines the idea, while another individual—an artisan or technician—realizes it physically. The legal dispute between Italian artist Maurizio Cattelan and French sculptor Daniel Druet exemplifies this. Druet, who created several of Cattelan's famous sculptures, including the iconic kneeling Hitler, sued Cattelan in 2021, claiming authorship and seeking recognition and compensation.[8] The court ruled in Cattelan's favor, upholding the principle that in conceptual art, the idea's creator is the author and copyright holder, even when others execute the physical work. Cattelan's role thus resembles an art director's, defining the conceptual framework while leaving physical execution to others. The Renaissance workshop model often involved direct master involvement, even if mediated by others' hands. The master's signature guaranteed quality and authenticity. Contrastingly, Cattelan's approach more sharply divides conceptualization and execution, often excluding the artist's hand from physical creation. This shift emphasizes ideas over craftsmanship, reflecting changing values in contemporary art.

In AI-assisted art, the lines between creator, tool, and collaborator are increasingly blurred. The key question is: At what point does AI transition from tool to collaborator

or creator? This transition likely depends on both the AI system's capabilities and how the human artist chooses to use it. A passive user who simply follows the AI's instructions to realize a work or artifact essentially attributes greater creative agency to the system. Conversely, an expert user with technical skills and clear vision uses AI to enhance their work while maintaining creative control, relegating AI to automating technical tasks without influencing artistic vision. This mirrors the difference between an expert photographer manually adjusting camera settings and a novice relying entirely on automatic modes.

We can draw parallels between AI-generated artworks and the artist-technician relationship, with output control varying based on the artist's involvement and reliance on the AI "artisan." However, a key difference lies in the usage rights: while AI-generated works can often be freely used by their creators, it is debated if they should be considered the creators' own works. For instance, when a user utilizes a paid version of Midjourney to generate an image, they receive a license to use the image for personal and commercial purposes. In this sense, they "own" the image but do not obtain exclusive copyright. This situation implies that a third party could potentially use the same image for commercial purposes without infringing on copyright laws (though it would breach community guidelines), as there is no exclusive copyright on the image itself.[9]

Who owns the copyright when an AI system generates a piece of work?[10] A recent case that highlights this issue is the U.S. Copyright Office's initial decision to revoke the copyright of the graphic novel "Zarya of the Dawn," by Kris Kashtanova, after discovering that the images were generated using Midjourney. The decision was based on the premise that there must be "substantial human involvement in the creative process" for a work to qualify for copyright protection. However, partial rights were later granted because the arrangement of the images and the text of the story were the product of Kashtanova's own creative efforts.[11]

European Union law distinguishes between different stages of the creative process: conception, execution, and "redaction" (that is editing, modification in post-production, refinement). Even when using AI, humans might not control the execution but still have authority over the conception and redaction stages, potentially allowing them to claim copyright. However, proving that a human has exerted sufficient creative control and intellectual effort to claim authorship remains a challenge.[12]

Since AI-system cannot copyright their own work, there are two prevailing theories for assigning copyright ownership of AI-generated work. The first theory considers AI as if

it were an employee working for an individual or organization, such as a company. Under this view, the copyright for AI-generated work would belong to the entity responsible for creating or operating the AI. This could be either the individual programmer who developed the AI or, more likely, the company that employs the programmer and owns the AI system. The second theory treats AI as a consumer product, in which case the end-user or customer who uses the AI tool would hold the copyright. [13] While a programmer is indeed responsible for creating the AI system's creative capabilities, this alone may not be sufficient to establish ownership rights, particularly in the case of generative AI. In these advanced AI models, the programmer merely creates the *potential* for output generation, rather than directly producing the final work. Similarly, it would be inappropriate to assign copyright of a painting to the artist's teacher rather than to the artist themselves. The teacher, like the AI programmer, provides the tools and knowledge, but does not create the final work. This situation is fundamentally different from earlier, rule-based AI systems where the programmer's role was more direct, since there was no involvement of intermediate users, and had a significant impact on the output.

The Goddess of Chance: The (perceived) Autonomy of Randomness

In the question of the relationship between tools and agency, a difference in attitude emerges between professionals and the general public. On one hand, professionals aspire to use AI as a tool, aiming to maintain detailed control over the output and preserve the possibility of customization and fine-tuning. From this perspective, "generative AI" tools would be nothing more than advanced forms of digital processing, comparable to software like Photoshop.

Systems based on diffusion models, like MidJourney, are significantly more complex. These models learn to map textual descriptions to visual elements during their training phase. The process begins by adding noise to images, creating corrupted versions. The model then learns to reverse this noise, gradually refining the corrupted images back to their original state. In the generation phase, when given a text prompt, the model starts with random noise. It then progressively reduces this noise, guided by its learned text-image associations, to create an image matching the description. This "denoising" process transforms abstract, compressed information into detailed, coherent outputs. The final image is a blend of the model's structured knowledge and random elements.

While random processes enable text-to-image models to create diverse outputs, users can exert precise control over the details through various methods. For example, they can provide specific feedback to iteratively refine the generated images, gradually steering the output toward their desired result. Another is through conditional inputs, where users provide detailed textual description, or mix text with sketches and reference images. By means of latent space manipulation users can interact with the latent space representation of the image by tweaking latent vectors. A user can control all levels of the generation, from very global attributes to very fine details, thus controlling different layers of generations.[14]

Many are fascinated by the patterns generated by a *kaleidoscope*, which produces regular aesthetic structures through a system of mirrors, resulting from random processes that arrange colored fragments inside a tube. Subjectively, we are inclined to attribute creativity or even a kind of autonomy to unpredictable and random outcomes. We tend to perceive something as creative and autonomous when we cannot rigidly predict its behavior. Conversely, what is predictable and determinable appears neighter autonomous nor creative to us.

Also the public's fascination with AI similarly stems from its unpredictability. If AI systems were perfectly controllable and customizable - as professionals might prefer - they might lose their appeal as "AI". We tend to call systems "AI" when they surprise us with unexpected outputs. This unpredictability comes from the use of stochastic processes in both the learning and generation phases of AI. These random elements prevent the system from being entirely deterministic and predictable.

However, not everything that is the result of chance and unpredictability looks creative or autonomous; disordered and meaningless chaos does not. Instead, it's the unpredictability imbued with meaning that captures our imagination as potentially creative or autonomous. The geometric shapes in a kaleidoscope represent a blend of randomness and structure, producing distinct and unique patterns with every turn.

This brings us back to the classic experiments of historical avant-gardes with chance and randomness. Movements like Dadaism and Surrealism, as well as subsequent experiments of Neo-Dada, for example in John Cage's aleatoric music (which was inspired by Marcel Duchamp's works on random music generation), were fascinated by the possibility of creation that escaped authorial control. Chance enables the creation of

works that look less artificial—that is, less like products of human artifice—by mimicking the spontaneous phenomena of physical nature.[15] As Cage said: "My *intention* is to let things be themselves".[16] A line of continuity can be traced between these experimentations and subsequent explorations of generative art and computer art, where the unpredictable element becomes an essential component for producing something that has the appearance of creativity that goes beyond the hard-wired instructions inserted by the programmer.

It's interesting to note the conceptual origin of terms that are apparently opposite to each other, such as "autonomy" and "automaton".[17] The ancient Greek noun *automatismos* is defined as "that which happens of itself, by chance", the verb *automatizo* means to "act of oneself, act offhand or unadvisedly", "[to] act spontaneously", [to] happen of themselves, casually". Similarly, "Automatia" was an early name for Fortuna, the goddess of chance[18]. Aristotle, in the second book of his "Physics," introduces the term "automaton" to describe a set of phenomena that includes "tyche" (fortune). He portrays these phenomena as forces that disrupt the teleological order, essentially embodying random or chance events in nature[19]. Aristotele stresses the role of the accidental also in the aesthetic domain of the tragedy. In "Poetics" (Book IX), he emphasizes how unexpected events can become especially meaningful in storytelling, but that coincidences in tragedy are most impactful when they seem to have an underlying design or purpose (an "air of design"), despite their apparent randomness.[20]

In essence, it's not the inherent randomness of spontaneous phenomena that surprises us; if it were the case, we'd find every random sequence of coin tosses or dice rolls surprising and creative. Rather, what captivates us is the apparent significance of these occurrences – when chance events seem to possess an *air of design*. This is why we aesthetically admire the patterns in a kaleidoscope: although they result from the random and spontaneous arrangement of colored fragments, the symmetry created by the mirrors gives them a sense of structure.

Although Aristotle suggests that chance events in physics lack discernible causes, hinting at a form of autonomy, these phenomena still emerge from natural processes and fall short of genuine autonomy in the fullest sense. True autonomy goes beyond unpredictability and accidental spontaneity, it requires agency and the ability to *initiate* action. The truly autonomous subject is, so to speak, left to venture alone into the world, guided by its system of motivations and knowledge, which is partly shaped by contingent encounters throughout its experiences. In more philosophical terms, using a

concept from Saint Augustine, an autonomous being must possess an *initium*—a beginning that is not causally determined. This idea has been revisited in contemporary times by Hannah Arendt in her book *The Human Condition*.[21] For Arendt, the essence of human freedom lies in the subject's ability to "enter the world" without pre-determination, equipped with the capacity to initiate something new.

In the context of AI, we are still far from creating entities that we may "let go into the world", giving them such autonomy as to make the machine's action completely unpredictable. This is partly because we would need to equip them with objectives, a motivational system, impulses, instincts, and "needs" whose (dis)satisfaction would drive them to act. Otherwise, even the most virtuous AI-artist, free to choose, might end up happily sitting idle. Instead, we grant these entities degrees of freedom in the form of predetermined doses of randomness.

In contemporary art, there is an interesting parallel with practices involving entities that possess their own natural agency, such as animals or plants. Examples include artworks were growth chambers encouraging specific types of vegetation to bloom, installations cultivating bacteria or fungi, or performances where artists interact with animals. These artists "use" nature—complex, spontaneous mechanisms guided by intrinsic principles rather than artistic design. In this way, they free natural elements within their work, allowing them to unfold with minimal intervention. These "agentic tools" occupy a middle ground between mere instruments and fully autonomous agents.

Artificial agentic tools also appear in artistic practices. For example, artist Mario Klingemann developed his project BOTTO as a decentralized autonomous artist that creates artwork using AI and community input. Launched in 2021, BOTTO produces about 350 new pieces weekly, voted on by a community of 5,000. These votes influence future creations by giving an aesthetic ranking. Klingemann sees his role as a guardian, initially guiding BOTTO but allowing it to gain independence over time. According to the artist, this approach resembles releasing a child into the world, trusting it will continue as its creator intended, much like parents do, but without having control over it.[22]

Artificial Author and Authorial Intentionality

Let's return to a similar scenario with which chapter 3 opened: imagine you've just finished reading a novel that has left you spellbound. The prose is mesmerizing, the characters lifelike, the plot rich and emotionally resonant. You feel a deep connection

with the author, admiring her emotional depth and understanding of human nature. Then you discover that the work was actually produced with an AI (or more precisely: by humans who used an AI). How would this make you feel? Would your admiration wane, to be replaced with a sense of disillusionment or even betrayal? Would the story lose its vibrancy? The characters in the novel feel less real?

AI's emergence is reshaping how the public and users perceive authorship, a concept related to but distinct from copyright, as it involves the perception of the source of a creative act to which we attribute both the origin and value of a work. The current debate is marked by heated opposition, where discussions often confuse the issue of the quality of AI-generated works with the question of their artificial nature. When these works are deemed meaningless or "not art," it is often unclear whether this judgment arises because they are aesthetically poor or simply because they are artificial.

The more traditional notion of authorship holds that part of the aesthetic appreciation we have for a work of art, a song, or a novel depends essentially on our awareness that behind the work there is a creative intention, along with the inferences we can draw from what the author intended to say. In other words, we take for granted that perceiving the mind behind a work of art is a fundamental component of our aesthetic engagement. It follows that we may not truly appreciate a work knowing it is the product of a machine lacking authorial intent, experience, or even consciousness, but merely "instructed" by humans to generate such products.

The question of authorship has been a focal point in philosophical and literary discourse, particularly in the context of structuralism and its successor, post-structuralism, according to which it is not possible to attribute a privileged and unique source of meaning to an author of a work of art. Roland Barthes famously declared the "death of the author" in his 1967 essay.[23] He argued that the author's intentions and biographical context should not dictate the interpretation of a work, since every text is the product of a complex web of influences, a "tissue of citations" and traditions, the recombination of a whole cultural past and social context. A text is the result of other texts and authors that speak through the pen of the alleged "autonomous" writer. Even before structuralism and post-structuralism, according to the New Criticism it was important to avoid the "intentional fallacy", a term coined by W.K. Wimsatt and Monroe Beardsley in 1946.[24] This principle challenges the practice of interpreting and evaluating a work primarily based on the author's intention, arguing instead for a focus on the work's content and the reader's experience. According to the New Critics, the author's intention is neither available nor desirable as a standard for judging the

success of a work of literary art. Instead, New Criticism promoted an analytical methodology that emphasized the text's complexity, unity, and the interplay of its constituent elements.

This tradition can also be linked to earlier avant-garde art experiments that explored artistic creation through reducing or even neutralizing the artist's control. As we saw in the previous paragraph, avant-garde movements like Dadaism and Surrealism used techniques such as chance and automatism to encourage spontaneous and collective creativity, thereby lessening the artist's role. For example, the Surrealists' endeavor to emulate a "machine-like" state was evident in practices like automatic writing. However, while avant-garde experimentation and automatically generated art emphasized liberation from human decision-making and control by means of stochastic processes or mechanic procedures (that nevertheless, one should not forget, still adhered to the creator's initial intent), contemporary AI-generated content introduces a novel form of autonomy with its own control and decision-making capabilities: human authoriality seems replaced by another form of authoriality that *mimics* the human one. Moreover, theories about the "death of the author", though philosophically compelling, have not actually truly manifested in public attitudes. Humans perceive intentionality in everyday life and, as a result, we continue to think in terms of authorship when encountering human-made cultural products.

In understanding the mechanism of authorship attribution, we could define two conceptual and psychological *thresholds*, which I would call the "threshold of instrumentality" and the "threshold of authorial relevance":

a) The first threshold, discussed in the previous pages, concerns the general question of the boundary between agency and its tools—specifically, where we draw the line between viewing a system as possessing autonomy or as a mere extension of the user. This threshold addresses the question of "where" the author is situated. The intermediate cases we have examined (such as assistants, collaborators, etc.) suggest that this threshold is not a distinct line but rather consists of a spectrum of intermediate cases.

b) The second threshold, the threshold of authorial relevance, focuses on determining in which contexts the presence of an author is relevant. As we discussed in Chapter 3, there are certain areas where the perception of an author behind a work seems relatively unimportant for aesthetic appreciation. While agency and intentionality are

significant in some forms of cultural production, they are not necessarily crucial in others.

On one hand, we can aesthetically engage with phenomena that lack a "mind," as they are not the product of human activity—like landscapes, flowers, or other natural structures, although, for some, even these might be considered the result of divine intentional creation. In cases where authorial intention is irrelevant to our appreciation, we tend to focus primarily on the formal, aesthetic features of the work. For example, in a beautifully designed piece of furniture or a modern architectural structure, we might concentrate on the form, the lines, the materials used, and the overall visual harmony. Similarly, we appreciate harmonious and low-fi background music for its pleasant and relaxing qualities.

In the 1950s experiment with algorithmically generated Mondrians by A.N. Noll, it was observed that the artificial images were favored over the originals for their formal qualities. This preference can be attributed to an exclusive focus on their abstract, non-figurative style, which lacks direct symbolic meaning for the naïve observer, whose judgment ignores the link between the image and the artist's original intentions. In a way, a purely formalist view of abstract art, one that disregards the author's sense-making, risks diminishing works like those pseudo-Mondrians to simple decorative patterns. Similarly, if we stand in front of a Pollock drip painting, we are drawn into a web of colors and patterns. But we also seek to understand the passion and the turmoil that the artist might have felt while creating this piece. The pursuit of unraveling the artist's intent adds a layer of depth to our aesthetic experience that an AI-generated Pollock, even thought formally impeccable, would not be able to offer.

A possible consequence of these considerations might be that the threshold of authorial relevance could be a demarcating criterion between "true" art, rich in meaning and relevant to the individual's subjectivity, and purely decorative, entertaining art. AI might thus find its niche in art forms where the "surface" aspect is paramount, and the presence of an author is not crucial for our enjoyment. This includes areas like background music, decorative patterns, industrial design, and formulaic narrative texts, among others. This inevitably raises the question of where to draw the line regarding the necessity of an author. When does the presence of a recognizable mind behind a work become essential for our aesthetic appreciation, and when can we do without it? The key point here is that neither of the two thresholds are fixed; their criteria for demarcation can change based on how we perceive and attribute instrumentality or intentionality. Most importantly, for works where we consider authorship fundamental,

we must consider whether this threshold remains valid or shifts, allowing us to view works where authorial perception was once crucial with a new authorless perspective.

Based on the discussion above, we can summarize several theories regarding the issue of AI authorship. The first two represent opposite ends of a spectrum: the first pole is

 1) The human-centric view, which considers the human author as the sole and exclusive source of all creation, with AI functioning merely as a sophisticated tool. Even when the human role is reduced to curating the machine's output, it is still the human who completes the final work. This includes concepts such as: 1.1) the "author as selector," where a person uses an AI system to generate a variety of images, sketches, designs, or texts and then curates and selects from these outputs what best aligns with their preference. Alternatively, 1.2) there is the notion of the human author as an instruction-giving "prompt-engineer," which involves using AI systems guided by linguistic "prompts," representing a newer form of indirect authorship (see Chapter 9).

2) The second pole sees AI as a full author. In this (potentially future) scenario, artificial intentionality or a semblance of mind would eventually be attributed to and recognized in AI-generated works. As AI technologies become more advanced, we may increasingly view their output as the work of entities with their own agency and intentionality. We might even imagine a context where AI is "free" to determine its own creative intentions and motivations, exploring and creating based on autonomous decisions.

3) An alternative interpretation, which draws from post-structuralism and critiques of the "intentional fallacy", views the authorship of AI-generated works as the result of a blending of sources, texts, and materials on which the systems have been trained. "Remixed authoriality" in the context of AI art suggests that works are seen as amalgamations of various influences, rather than as products of a single creative mind. This view aligns with the idea that all cultural artifacts are inherently "post-productive," meaning they are reconfigurations of pre-existing materials, challenging the traditional, romanticized notion of authorship as a unique creative expression of an individual. In this perspective, the author's role is similar to that of a curator who brings together diverse cultural elements but is not the sole source of authorship. Instead, the author becomes a conduit of a collective authorship embedded in human cultural history. This approach is especially relevant in the discussion around "remix culture", where creation involves recontextualizing, quoting, and repurposing of existing works.[25] Within the "Remixed authoriality" framework, AI systems draw from extensive databases encompassing various domains of human culture and serve as a medium through which

a wide array of human expressions, ideas, and cultural artifacts are processed and reinterpreted. The resulting creations are not just the products of programming by human creators but also reflections of collective human intentionality. Therefore, the outputs of these systems can be seen as manifestations of *collective authoriality*, filtered and transformed by the artificial system.

These three conceptions of authorship (the human, the machine, and the collective authorship) reflect an underlying conceptual need to identify an inspirational source that then takes shape in the intentional construction of a work. In this sense, these theories differentiate between the source of inspiration and the intentional process of a work's construction, defining various and shifting dependencies between the creator and its different "executors." Historically, entities such as "God," "the Muses," or more recently, our "cultural memory," have been seen as the primary sources of inspiration, with authors acting as channels for these higher forces. Similarly, although AI is initially designed to assist authors, its increasing complexity and influence allow it to generate content, provide inspiration, and shape creative processes. Consequently, AI might evolve from being viewed as a tool to becoming a genuine source of inspiration, with human intentionality acting as the instrumental executor of this inspiration. The individual would increasingly serve as an intermediary or facilitator, functioning as a tool for a distributed authorial intention that permeates our cultural archives and the technological means of their expression.

4) A further possibility is that AI-generated works compel us to abandon any inferences about authorial intention. In this scenario, we cease to attribute any mind behind the AI artwork, limiting ourselves to a purely formal appreciation, akin to our response to decorative patterns or design products that captivate us primarily for their superficial appearance. If the focus may move away from the idea of authorial intentionality, the primary concern would be if a work resonates with us on a personal level regardless of any hypothesis concerning the creator's identity, whether human or machine. This shift would represent a significant change in how we engage with creative works, shifting the center of our attention on the direct exposure to formal and aesthetic qualities of the work, avoiding questions about its origins. Alternatively, since we do not recognize any authoriality in AI-generated works, we could altogether avoid them, considering them "soulless" and therefore unworthy of our attention compared to true human works. Consequently, it would be significantly impactful for us to know with certainty whether the music we are listening to or the novel we are reading was produced by a human or a machine, as we might suspend our judgment on authorship and thus any aesthetic engagement only in the case of machine-produced works.

The idea that in the future there could be two distinct approaches to authorship—one for human-made products and another that does not attribute authorship to artificial products—is, however, naïve for at least two reasons. First, it will not always be possible to determine whether a cultural artifact or product—be it a text, image, sound, or piece of music—was created by a human or predominantly generated by AI. Second, our relationship with new forms of technological and cultural production could radically alter how we perceive and think about *any* kind of authorship, including traditional human authorship.

Imagine a future where it becomes increasingly challenging to determine whether a painting, a musical composition, or a written text is the product of human intelligence or artificial process. In such a scenario, the significant shift might not be in how we perceive the authorship of machines, but rather in our overall understanding of authorship. It is conceivable that our expectations and inferences about authorial intentions may weaken and diminish due to the persistent doubt over whether there is any author at all behind what we are observing. The constant uncertainty about the origin of these works might lead us to approach them with a different mindset, one less concerned with discerning the creator's identity and more focused on the work itself, independent of the underlying creative intentions. This shift could fundamentally alter how we interact with and appreciate artistic and creative works.

This "post-artificial" stance, as articulated by H. Bajohr[26], foreshadows a radical shift in our approach to understanding and interacting with texts or other artifacts. The pivotal question concerns how we read a text or listen to a song when we can no longer be certain whether it was written by an AI or a human. On one side, as we discussed, this situation could open the door to the humanization of machines, suggesting that we might start to see AI as more than just tools or mechanical aids. On the other side, it also prompts a reevaluation of the human creative process, recognizing the "mechanical" aspects inherent in our own creativity and intentionality.

A "post-artificial" scenario could, for some, be seen as the practical realization of the "death of the author" theory proposed by post-structuralists. This scenario eliminates the concept of authorship, implying a lack of direct dialogue between the audience and the author, since the author is no longer present. However, this does not mean that the internal dialogue between the audience and the work ceases to exist; the work itself can express a form of immanent "authorial voice". Perceiving a specific intentionality in the text does not necessarily imply making assumptions about the actual process that

produced that text (see Chapter 3). In this regard, narrative theory traditionally distinguished between *real* and *implied* authors.[27] While the former is the actual writer of the text, the latter is the voice grounded in the text and expressed by its content and style. The implied author thus becomes a reader-created construct that is different from who (or what) the actual creator is: when we read a text, we imagine the writer, his thoughts, and his personality emerging from his choice of words, expressions, and sentences. Therefore, though we may know that a text has been artificially generated, we could still engage with the implied author expressed in the text, immersing ourselves in what he or she has to say. Similarly, in other artistic expressions as well, the crucial factor may be the ability of the artifact to "express" intentionality and motives, effectively allowing the construction of an authorship that emerges from the work, over and above the actual source that produced it.[28] Therefore, we might instead limit ourselves to attributing an "implied" author, assuming a stance in which we relate to the work *as if* there were an actual intentionality, suspending our judgment about the presence of a "real" (that is, human) author.

Alternatively, in the absence of an author, the viewer/listener might put themselves in the perspective of a potential author, mentally simulating their presence. In this case, the implied authoriality would become an actively *imagined authoriality*, similar to what happens in the imaginative play we engage in when observing random, inanimate forms (lines on a rock, cracks on a wall, cloud formations) and assuming that the patterns we discern in them are the result of intentional design.

Where Does "Effort" Go?

"This song sucks. […] Songs arise out of suffering, by which I mean they are predicated upon the complex, internal human struggle of creation and, well, as far as I know, algorithms don't feel. Data doesn't suffer. […] Writing a good song is not mimicry, or replication, or pastiche, it is the opposite. It is an act of self-murder that destroys all one has strived to produce in the past. […] It's a blood and guts business, here at my desk, that requires something of me to initiate the new and fresh idea."
This is the 2023 passionate response by songwriter Nick Cave, who runs a blog called *The Red Hand Files*, to a particularly enthusiastic fan who wanted to pay tribute to the singer with a song generated by ChatGPT "in Nick Cave's style."

A critical viewpoint raised in the debate on AI-generated art is that the value of such systems' outputs is questionable due to their quick, automated and apparently effortless

processes, among other reasons, as in the example of a "Nick Cave"-like song that just emulates the stereotype features of the songwriter's texts. Those products seem to lack human *effort*, there is no sign of any kind of creative struggle, which may be defined as overcoming some material, technical but also cultural and mental barrier. "Effort" also means being always plagued with the uncertainty of the outcome and the possibility of failure. An artist's creative process is more like an uncertain exploration, where the artist moves through various ideas and possibilities without a fixed path. On the other side, what these machines do appears too smooth, mechanical, and pre-determined.[29] This critique also echoes certain debates that animated the 20th century, when the public and some critics questioned the apparent "ease" of certain experimental avant-garde art forms, such as ready-mades, abstract minimalism, - think of Fontana cut-canvases – conceptual art, and simple performative acts. The criticism was mostly directed to the perceived simplicity of the artists' creative "gesture", which was not considered sufficient to confer value on their work.

If (perceived) effort determines our aesthetic judgment, do we look, listen or read the work with different eyes according on how much "suffering" there is behind? Empirical evidence seems to confirm the tendency – defined as "effort heuristics" – to use effort as a proxy of aesthetic value.[30] The question that consequently arises is: could this be an element of our suspicious attitude toward AI-generated artworks? Or, alternatively, could machines suffer (make effort) or at least, *show* effort?

On this point, we might observe a historically ambivalent attitude toward "effort" in craftmanship: the degree of an artist's or craftsman's skill could reveal – as said – what Leonardo da Vinci called "ostinato rigore" (stubborn rigour). But from another perspective, the wonder and admiration of the artist's product might even be inversely relates to the effort exercised in creative process. Skilled artists or craftsmen can produce artifacts with less effort compared to novices. According to a famous quote attributed to Michelangelo: "If people knew how hard I had to work to gain my mastery, it would not seem so wonderful at all". That is: the talented artist is able to show *geniality* or (even divine) inspiration *not in his effort,* but in his *ease* in accomplishing what others cannot do, or in manifesting "sprezzatura", namely the apparent effortlessness in his craft. The Latin saying "Ars est celare artem," which translates to "Art is to conceal the art," emphasizes the idea that true and valuable art often hides the effort put into its creation. This principle suggests that the most impressive art appears effortless, even though it may require immense skill and labor.[31] The artist's struggle, experimentation, and refinement are hidden behind the final product, which appears

natural and spontaneous. Contrary to that, excessive effort can, in a way, indicate a lack of experience or inspiration.

The Romantic era focused on the artist's internal struggle, shifting emphasis from technical proficiency to emotional and intellectual effort. This anticipated the later shift in Avant-Garde movements that valued the novelty of the idea (and the effort put to come up to it) as more important than the length of time and effort in crafting a work. Lucio Fontana's cut canvases serve as a prime example of this shift. Fontana's straightforward yet innovative act of slicing through the canvas redefined the concept of effort, moving away from labor-intensive techniques towards an emphasis on conceptual depth. Due to its non-obvious nature, these kinds of contemporary artworks always risked being perceived as irrelevant or less engaging. Consequently, it became imperative for artists, critics and art gallery curators to clearly articulate how a piece of conceptual art embodies effort and skill on a more abstract, spiritual and mental, level.

Similarly, common opinion often sees AI art as "too easy" and lacking effort: however, this view could be countered at least by two perspectives. The first aspect is that, beyond the most amateur uses, the work of artists who work with AI systems is not at all "simple" or automatic but requires both a deep knowledge of the medium and the technologies used, as well as a complex and detailed practice in the realization of the final work. For example, creating a sufficiently sophisticated prompt to realize the artist's precise creative idea is often an intense and demanding job that has nothing automatic or simple about it. The second aspect concerns the question about the possibility to conceive a "machine effort" or an "artificial struggle" in the generation of content by AI-systems. One reason for this argument is based on the extensive effort in developing AI technologies and the human contributions implicit in the training dataset of these systems. The outputs of these technological systems arise from the amalgamation of varied influences and content from different times, as those systems learn from the works of past artists, such that we might here speak of a *distributed* aesthetic effort, akin to what we previously called a "collective authoriality", that allows us to value the whole sum of past influences, individual contributions, and also technological evolution that led to a specific artifact or artwork. The distributed nature of AI's development and learning process means that its effort is a collective one, spanning many individuals and technological advancements. The ease in the generation of content is just apparent, a kind of "artificial sprezzatura", where the seemingly automaticity in AI-production masks the vast cultural knowledge on which the AI has been trained, the complex computational work, not to mention the significant use of material and energy resources that those systems employ.

Notes

1 Dean, M. (2019). Artist Mario Klingemann on Artificial Intelligence, Technology and our Future (interview with Mario Klingemann), Sotheby's, https://www.sothebys.com/en/articles/artist-mario-klingemann-on-artificial-intelligence-art-tech-and-our-future.

2 Arielli, E. (2018). Taste and the algorithm, *Studi di Estetica*, 12, 3, 77-97.

3 Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2024). Generative AI can harm learning. SSRN. https://ssrn.com/abstract=4895486.

4 Finn, E. (2017). *What algorithms want. I imagination in the age of computing*. MIT Press.

5 Morris, M. R., Sohl-Dickstein, J. N., Fiedel, N., Warkentin, T. B., Dafoe, A., Faust, A., Farabet, C., & Legg, S. (2023). Levels of AGI for operationalizing progress on the path to AGI. *Proceedings of ICML 2024*. arXiv:2311.02462.

6 Subbulakshmi, N., Begum, S. A., Archana, B. L. G., & Chandru, R. (2024). AI-based emotion-aware music composition using Spotify app. In *2024 3rd International Conference on Sentiment Analysis and Deep Learning (ICSADL)* (pp. 59–63). https://doi.org/10.1109/ICSADL61749.2024.00015.

7 Butlin, P., Long, R., Elmoznino, E., Bengio, Y., Schwitzgebel, E., et al. (2023). Consciousness in artificial intelligence: Insights from the science of consciousness. *CoRR*. https://arxiv.org/abs/2308.08708.

8 Lydiate, H. (2022). The unasked question. *Art Monthly*, (458), 44–45.

9 Misthal, M. P. (2024). AI-generated works dilemma: Balancing AI terms of service with contractual obligations. *The Legal Intelligencer*. https://www.law.com/thelegalintelligencer/2024/04/16/ai-generated-works-dilemma-balancing-ai-terms-of-service-with-contractual-obligations; see Midjourney's Term of Service as of 9/1/2024: https://docs.midjourney.com/docs/terms-of-service

10 Caldwell, M. (2023). What is an "author"? Copyright authorship of AI art through a philosophical lens. *Houston Law Review, 61*(2), 411–450.

11 Edwards, B. (2023, February 23). AI-generated comic artwork loses US copyright protection. *Ars Technica*. https://arstechnica.com/information-

technology/2023/02/us-copyright-office-withdraws-copyright-for-ai-generated-comic-artwork/.

[12] Militsyna, K. (2023). Human creative contribution to AI-based output – One just can('t) get enough. *GRUR International, 72*(10), 939–949. https://doi.org/10.1093/grurint/ikad075.

[13] Naqvi, Z. (2020). Artificial intelligence, copyright, and copyright infringement. *Marquette Intellectual Property Law Review, 24*(1), 15–51.

[14] Pathania, K. (2024). Enhancing conditional image generation with explainable latent space manipulation. arXiv:2408.16232.

[15] Mann, P. (1990). Symmetry, chance, biomorphism: a comparison of the visual art and poetry of Hans Arp's Dada Period (1916–1924). *Word & Image*, *6*(1), 82–99. https://doi.org/10.1080/02666286.1990.10435422; Jaeger, P. (2013). *John Cage and Buddhist ecopoetics: John Cage and the performance of nature*. Bloomsbury Academic.

[16] Cage, J., & Tudor, D. (1959). *Indeterminacy: New aspect of form in instrumental and electronic music* [Recorded reading]. Folkways Records.

[17] Autonomy, derived from "autos" (self) and "nomos" (law), implies self-government and self-directed action; *automaton* derives from "autos" (self) and "maton" (acting), often used to describe entities that move on their own or events that seem to happen without purpose or intention.

[18] Chiodo, S. (2022). Human autonomy, technological automation (and reverse), *AI & Society, 37*(1), 39–48, doi: 10.1007/s00146-021-01149-5.

[19] "But is it not equally strange that, however freely men admit that every kind of luck [*tyché*] and everything that 'happens accidentally' [*automaton*] can really be assigned to some definite cause, still, while accepting this venerable argument for the elimination of chance from their thoughts, they nevertheless invariably distinguish, in fact, between things that do, and things that do not, depend upon chance [*automaton*] or luck [*tyché*]? (196a12–17)" Aristotle. (1929). *Physics* (P. H. Wicksteed & F. M. Cornford, Trans.). In *Aristotle IV*. Loeb Classical Library. Harvard University Press.

[20] "Such an effect is best produced when the events come on us by surprise […]. The tragic wonder will then be greater than if they happened of themselves or by accident ("τοῦ αὐτομάτου καὶ τῆς τύχης", "tou automatou kai tēs tuchēs", translates also to "by chance and fortune" or "by accident and luck."); for even coincidences are most striking when they have an air of design." (Aristotle, Poetics, Book IX).

[21] Arendt, H. (1958). *The human condition*. University of Chicago Press.

22 Nair, S. (2022). German artist Mario Klingemann on his creation 'Botto', an NFT revolution. https://www.stirworld.com/see-features-german-artist-mario-klingemann-on-his-creation-botto-an-nft-revolution.

23 Barthes, R. (1967). The death of the author. In S. Heath (Trans.), *Image, music, text* (pp. 142–148). Hill and Wang (1977).

24 Wimsatt, W. K., & Beardsley, M. C. (1946). The intentional fallacy. *Sewanee Review, 54*, 468–488. Revised and republished in *The Verbal Icon: Studies in the Meaning of Poetry* (pp. 3–18). University of Kentucky Press (1954).

25 Lessig, L. (2008). *Remix: Making art and commerce thrive in the hybrid economy*. Penguin Press; Navas, E. (2022). *The rise of metacreativity: AI aesthetics after remix*. Routledge.

26 Bajohr, H. (2023). Artificial and post-artificial texts: On machine learning and the reading expectations towards literary and non-literary writing. *Basel Media Culture and Cultural Techniques Working Papers*, 2023.007. https://doi.org/10.12685/bmcct.2023.007.

27 Booth, W. C. (1961). *The rhetoric of fiction*. University of Chicago Press (New edition, 1983).

28 Pierosara, S. (2022). Narrative autonomy and artificial storytelling. *AI & Society*. https://doi.org/10.1007/s00146-022-01595-9

29 Chamberlain, R., Mullin, C., Scheerlinck, B., & Wagemans, J. (2018). Putting the art in artificial: Aesthetic responses to computer-generated art. *Psychology of Aesthetics, Creativity, and the Arts, 12*(2), 177–192.

30 Kruger, J., Wirtz, D., Van Boven, L., & Altermatt, T. W. (2004). The effort heuristic. *Journal of Experimental Social Psychology, 40*(1), 91–98. In this experiment, participants consistently showed a preference for a painting they believed took longer to create.

31 D'Angelo, P. (2018). *Sprezzatura: Concealing the effort of art from Aristotle to Duchamp*. Cambridge University Press.