In a world where machines do not stop at creativity, this work explores the creative possibilities of latent space, the mathematical space that underlies all recent generative algorithms.

In the context of the concepts of metamodernity and metacreativity, Benjamin Bertram illustrates that innovative technologies such as Stable Diffusion and ChatGPT function not only as tools, but also as creative partners that revolutionize the process of design. In this context, the differences between human and computer-aided creativity are highlighted through practical experiments. The question arises as to how variables can be manipulated in order to generate unexpected aesthetic experiences. To what extent does the factor of chance play a role in the creative process? What role do people play in the meta-creative process?

These questions lead to a profound reflection on the nature of creativity, which is increasingly moving away from traditional concepts. The work not only offers an insight into the technical aesthetics of AI-generated art, but also encourages a critical examination of the ethical and social implications of this new form of creation. With a view to future research perspectives and the dynamic interaction between man and machine, readers are invited to develop their own ideas.
to reflect on their own creative processes and expand the boundaries of what is possible.

To the website

Benjamin Bertram PASSIVE ILLUSTRATION Master documentation and theoretical work Summer semester 2024

Benjamin Bertram

Master's documentation and theory thesis
summer semester 2024

Hamburg University of Applied Sciences
Faculty of Design, Media and Information

First examiner: Prof. Peter Kabel Second examiner: Prof. Dr. Michaela Diener

ILLUSTRATION

Latent space as a meta-creative medium

[redacted]

ILLUSTRATION

Table of contents

1 Introduction

The rapid development of artificial intelligence *(AI)* has fundamentally changed numerous areas, including art and design. One of the most important advances is the rise of generative AI, in particular algorithms for text and image generation. These systems, such as **Midjourney**, **DALL-E** and **Stable Diffusion**, have opened up new possibilities for artists to generate and explore artificial content. **Latent space**, a mathematical construct underlying these generative models, serves as a medium through which new creative possibilities arise *(Kingma & Welling, 2019)*.

This thesis aims to explore latent space as a new creative medium, specifically using the example of **portraiture**. By using AI-generated art and examining the interaction between machine-generated creativity and human curation, the thesis will demonstrate how latent space can be **systematically** navigated to discover new artistic styles and forms of expression *(Akten, Fiebrink, & Grierson, 2020)*. The focus on portraiture serves as a lens through which the complexity of machine-driven creativity can be examined in comparison to traditional forms of human creativity.

1.1 Context and motivation

Since 2022, text-to-image algorithms such as DALL-E and Stable Diffusion have reshaped the landscape of digital art and illus- tration. These models can produce images that mimic the look of professional photography or hand-drawn illustrations, but do not fall under the traditional definition of either, as they are merely an emulation of the trained data material *(El-*
gammal et al., 2017). The images produced by these systems are computer-generated illustrations created by algorithms that interpret text input and transform it into coherent visual outputs. This raises important questions about the importance of human input in this process: What kind of creativity is present when a machine creates the image? And how does the human role change when machines do most of the work?

In this context, the concept of **curation** in
of machine-generated creativity. Designers are no longer the sole authors of the artwork, but control and refine the output produced by AI systems. This creative process is primarily a curating one - it manages parameters, selects favorable results
and directs the machine output from randomness to structure. This also means that creative people no longer actively participate in the creative process, but are passively creative: They are given the creative output of the machines and they respond by adapting or accepting it.

1.2 Objectives and structure of the work

The main aim of this research is to prove the hypothesis that latent space can function as a creative medium. Through experimental studies on portrait generation, this thesis aims to show how different stylistic approaches and artistic methods can be derived from working with AI image generators. Originality is evaluated not only on the basis of human input, but also on how much the machine autonomously contributes to the creative process *(McCormack & Samsel, 2024)*.

The work is divided into six chapters:

- Chapter 2 provides the **theoretical framework** and gives an overview of the concepts of AI, generative AI and the historical development of machine-generated art.
- Chapter 3 deepens the theoretical framework around creativity by **comparing** human creativity with machine-driven processes and examining the role of latent space in enabling novel artistic combinations.
- Chapter 4 presents the practical **experiments** and shows, how AI-generated images can deliver unexpected but aesthetically convincing results with regard to portraits.
- Chapter 5 concludes with reflections on the wider **implications** of AI-driven creativity for the future of art and design and provides an **outlook** for further studies.

1.3 Theoretical basics

This work focuses on the concept of latent space. This is a **multidimensional mathematical space** in which learned representations of data are stored. Generative AI uses the latent space to encode abstract features of the input data (such as color, shape and texture), which can be manipulated to generate new results. When navigating through the latent space, the user or
the user can explore unexplored areas of creative potential and achieve results that may be completely novel or even unexpected. The exploration of this space allows for a mix of intuitive artistic decisions and algorithmic combinatorics, revealing new ways of thinking about the intersection of technology and art.

In this context, this paper critically examines the evolving definition of creativity. Traditionally, creativity has been viewed as an inherent human trait associated with emotional depth, cultural expression and social interaction, but a challenge for AI *(Boden, 1998)*. However, the introduction of Artificial Creativity complicates this view. By comparing human creativity - which is often influenced by social, emotional and cultural factors - with the logic-driven processes of machine creativity, this thesis challenges conventional notions and proposes new frameworks for understanding the role of algorithms in the creative process *(Davis, 2021)*.

1.4 Relevance for design and illustration

This research is particularly relevant to design disciplines, especially **illustration**, where the impact of AI is already being felt. AI-generated images are increasingly being used in commercial design and art projects, and although most illustrators are not yet working with AI, **a quarter** are already experiencing difficult working conditions such as unrealistic pricing or fewer commissions *(Illustratoren Organisation e.V., 2023)*.

Another important topic is the images themselves: While generated images can mimic the look of traditional illus- trations, they are fundamentally different in their origin. They are the result of complex algorithms working within the constraints of a training dataset and not the direct result of the designer's hand. This raises fundamental questions of authenticity, originality and intellectual property rights, as the dataset is based on millions of images collected from the internet *(scraping)*.

Similar to the invention of photography in the late From the 19th century, 3D tools in the 90s, and the rise of creative coding and thus also creative design during the turn of the millennium, generative AI is a new tool that speaks a new language. Although it is getting better and better at **mimicking** existing media, it also has its own visual uniqueness - but what does this look like, and how can the creatives produce them not just by chance, but consciously?

These questions will guide the experimental work in the following chapters, in which AI-generated portraits will be evaluated in terms of their artistic value and the role of the designers in their creation.

Fig. 1: Impact of AI on the illustrative business.

2 What is artificial creativity?

According to AI researcher Demis Hassabis, AI is a process that transforms **unstructured information** or data into **useful, usable knowledge** *(Royal Academy of Arts, 2018)*.

It differs from human intelligence in its ability to perform tasks such as visual perception, speech recognition, decision-making and natural language processing. AI has evolved from early symbolic approaches to modern machine learning *(ML)* and deep learning systems. Today's AI landscape is dominated by so-called Narrow AI, developed for specific tasks, while *Artificial General Intelli- gence (AGI)* remains a future goal, where machines could potentially match or exceed the cognitive abilities of humans *(Turing, 1950)*.

The development of AI can be traced back to the middle of the 20th century, with pioneering contributions from figures such as Alan Turing and John McCarthy. McCarthy's coining of the term "artificial intelligence" in 1955 marked the beginning of systematic AI research *(Mc-Carthy, 1955)*. Early AI models relied on symbolic reasoning, but with the introduction of **neural networks**, which enabled more sophisticated pattern recognition and decision making, the field made dramatic progress *(LeCun, Bengio, & Hinton, 2015)*.

2.1 Humans vs. machines: a comparison

AI is often compared to humans in various benchmarks *(Callaway, 2024)*. The goal of AGI is not only to match human capabilities, but to surpass them - but what are they?

If we take a descriptive look at the human species and its invention, the machine, we can examine areas that are comparable, but also areas that are unique to each **system.**

2.1.1 People

Humans are a species of about 8 billion *(United Nations, 2022)* upright walking apes with brains that enable language and tool use to overcome physical limitations.

The human brain has 86 billion neurons and 100 trillion synapses, which enables an enormous processing capacity; however, parallel processing (or multitasking) proves to be inefficient *(Tamagnini, 2023)*. Our memory can store up to 2.5 petabytes, which corresponds to around 30 million hours of television *(Salk Institute, 2016)*. About half of the brain is responsible for visual processing; we can recognize 10 million colors and have a visual bandwidth of 10-100 Mb/s. The auditory and tactile system processes around 1 Mb/s, while smell and taste reach around 10 Kb/s *(Stiegler, 2023)*.

Neuroplasticity enables adaptations, empathy fosters emotional connections and pain processing takes place on several levels. The autonomic nervous system unconsciously regulates heart rate and breathing, while sleep supports memory consolidation.
Food provides energy, ageing and death shape the life cycle, and our body has self-healing powers that are supported by the immune system *(Schmidt & Lang, 2019)*.

Humans have unique language abilities as well as areas for symbolism and creative thinking *(Siegler et al., 2021)*. Thanks to emotional intelligence, we can understand and process complex feelings. Our species maintains a culture with around 7,000 languages, art and rituals that promote social cohesion *(Carey, 2008)*. Humans form complex societies characterized by cooperation and cultural values. Stable and effective groups can comprise up to 150 individuals; beyond that, subgroups are formed *(Dunbar, 1992)*.

Fig. 2: The bechmarks for o1, OpenAI's latest language model, are evaluated using standard tests for different academic subjects.

The idea of comparing humans and machines as "systems" in terms of creativity was also noted in Davis, 2021, p. 9.

Humans are thus a highly evolved, socially networked species of upright primates with a complex brain that enables language, tool use, symbolic thinking and culture to overcome physical and mental challenges, while evolutionary adaptability and empathy strengthen our social cohesion.

In this context, creativity is not only a human ability that dynamically adapts to new situations, but also a form of social communication. When people's language skills decline after a brain injury, some resort to the expressive form of paintings and drawings *(Donald, 2020)*.

In late modern societies, creativity increasingly functions as an economic factor. In his book "The Invention of Creativity", Reckwitz argues that modern societies need an economy that is independent of artists.
have developed **a creativity dispositive**: a contradictory Connection between the desire for creativity and the imperative to be creative. People want and should be creative. In this context, the new plays a decisive role by demanding and promoting the circulation of goods, technical inventions and political revolutions. The consequences of this dispositive of creativity include pressure to perform, discrepancies between creative achievement and success, distractions and an overextension of the aesthetic, which is also reflected in the term "lifestyle" *(Reckwitz, 2013)*.

While creativity was reserved for creative artists in the pre-modern era, the term has moved closer to that of capitalist innovation in late modernity and thus tends to become an attribute of entrepreneurial activity *(Manske & Schnell, 2018)*.

2.1.2 Machines

One of the most important inventions of mankind is the principle of the machine. Machines are devices or systems that use mechanical, electrical or computational components to perform certain tasks or functions. They range from simple mechanical tools such as levers or pulleys to complex systems such as computers or robots. Machines often automate tasks and make work easier, faster or more precise. They can be powered by various energy sources, e.g. human effort, electricity or motors *(Grunwald, 2024)*.

Machines equipped with AI are technical systems,

Fig. 3: Mechanism of Antikythera, ca. 65 BC.

that use complex **algorithms** to solve tasks that conventional machines cannot handle. They have immense computing power with billions of transistors and specialized hardware such as graphics processing *units (GPUs)* and neural processors *(NPUs)*, which enable them to process large amounts of data quickly. While the human brain can store around 2.5 petabytes, the storage capacity of machines is theoretically unlimited, especially through cloud storage and distributed systems *(Backblaze, 2023)*.

Machines are characterized by multitasking, as they can carry out several processes in parallel and efficiently, with data transfer rates of several terabits per second *(Tb/s)*. AI systems such as GPT-4 are capable of speech and can understand and generate natural language. *(OpenAI, 2023)*

In the following, the word "machine" is synonymous with the conglomerate of servers, computers, programs and data that are important for AI systems.

Although machines do not have emotions or emotional intelligence, they can use algorithms to analyze and respond to human emotions, known as affective computing *(Somers, 2019)*. Machines do not have neuroplasticity, but they can improve their performance and adapt to new data through machine learning and neural networks. Machines use electricity directly as primary energy without converting low-energy organic or non-organic solids. Ageing and "death" affect machines due to hardware faults and incompatible software, but these can be rectified through maintenance and updates. Their further development is currently taking place through human-controlled technological advances.

Machines are networked, cooperate via the Internet of Things *(IoT)* and use cloud computing. They have no culture, but can analyze cultural content and creatively create new works *(Gabsi, 2024)*.

AGI refers to a form of AI that can perform general cognitive tasks at a level that matches or surpasses human intelligence. In contrast to today's AI, which is often limited to certain applications *(e.g. speech recognition or image processing)*, AGI would be able to react flexibly to new tasks without having to be specially programmed for them. AGI is seen as a long-term goal in AI research *(Gobble, 2019)*.

A machine is therefore a man-made system that uses mechanical, electrical or computational components to automate tasks efficiently and precisely. Modern machines with AI can solve complex problems and adapt to new situations through machine learning.

adapt to new data without possessing emotions or consciousness.

While creativity also fulfills a social function in humans, this is irrelevant for machines - but artificial creativity has a relevant advantage when it comes to target functions: Hassabis refers to the AI program AlphaGo, which derives patterns and possible positions from all recorded variants of the board game GO. The program extrapolates creative new strategies instead of simply averaging weights *(Grassegger & Krogerus, 2023)*.
In terms of new solutions that are both useful and practical, artificial creativity can not only match human creativity, but even surpass it. And the target function is variable: for the machine, it doesn't matter whether the aim is to develop optimal moves, drive accident-free or produce aesthetically pleasing images.

2.2 The development of generative AI

The idea of creating art with the help of machines is not new. Historical examples include **automata** - mechanical devices from the 18th century that could imitate human movements and create drawings.
These automata, such as the famous Maillardet automaton, represent early experiments with mechanical creativity, but are basically rule-based systems. They worked with pre-programmed mechanics instead of autonomous decision-making.

Fig. 4: Automaton by Henri Maillardet 1805.

The advent of digital computers in the mid-20th century revolutionized creative expression. Pioneering works such as those of Harold Cohen and John Whitney used computerized systems to create generative art that combined randomness and algorithmic design. Whitney's algorithmic films from the 1960s are an example of how early computer graphics laid the foundation for today's generative AI systems *(Moritz, 1997)*.

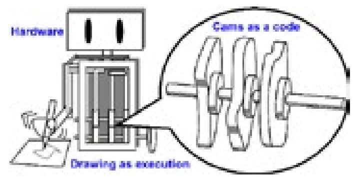


Fig. 5: How automata work.

With the development of Generative Adversarial Networks *(GANs)* by Ian Goodfellow in 2014, algorithms have not only made a significant leap in their "creative" capabilities, but have also become increasingly well known. GANs consist of two neural networks - a generator and a discriminator - that compete against each other and generate extremely realistic images from random noise. This breakthrough made it possible

machines to not only replicate, but also generate new, unseen images, and marked the birth of modern generative AI *(Goodfellow et al., 2014)*.
In addition to GANs, variational autoencoders *(VAEs)* and diffusion models have established themselves as important tools in generative AI. VAEs, introduced by Kingma and Welling in 2013, are probabilistic models that can learn the latent structure of data and generate new samples from the learned distribution *(Kingma & Welling, 2013)*. VAEs are particularly useful for creating smooth interpolations between different data points in latent space, making them powerful for image synthesis.

More recently, diffusion models have gained attention due to their stability and ability to produce high-quality images. These models, such as those used in Stable Diffusion and DALL-E, refine images incrementally, allowing for more controlled and precise image generation *(Ho et al., 2020)*.

2.3 Stable diffusion as a driver of innovation

The development of generative **open source AI tools** has fundamentally changed the landscape of digital creativity. The release of Stable Diffusion by Stability AI in 2022 ushered in a new era of accessibility for creatives and interested parties alike. Unlike previous proprietary models such as DALL-E, Stable Diffusion is open source and allows users to modify the model and adapt it to their specific artificial needs *(Stability AI, 2022)*. Although open-source image generators such as StyleGAN and VQGAN+CLIP already existed, it was the combination of easy access, controllability and excellent image quality that led to the widespread acceptance of Stable Diffusion. This accessibility quickly fostered the formation of an innovative community of creators and developers *(Vin- cent, 2022)*.

The emergence of platforms such as Civitai.com and the stable diffusion add-on **ControlNet** by Zhang et al. *(2023)*, developed from the open source community, has further advanced this democratization by allowing users to share fine-tuned models and gain greater control over output.
of AI-generated images. These tools provide a level of customization previously unavailable in proprietary AI systems and improve both the

user-friendliness as well as the creative potential of the imaging tools.

2.4 Computational Creativity: A paradigm shift

The development of generative AI has brought about a paradigm shift in our understanding of creativity. Traditionally, creativity has been seen as an exclusively human ability that is closely linked to emotional, cultural and social contexts. The advent of artificial intelligence is challenging this notion, as machines are now capable of producing highly creative works independently. **Computational creativity** differs from human creativity in that it follows programmed systems defined by data and thematic rules, rather than emotional or subjective judgment. According to computationalcreativity.net, computational creativity is

> *"... the study and simulation, by computational means, of behavior, natural and artificial, which would, if observed in humans, be deemed creative."*

This definition was dropped in 2013, but it seems to fit the concept better than the current definition from computationalcreativity.net *(Jordanous, 2014)*

Computational creativity can help us to understand our own human creativity, as we can judge whether something is creative or not - although a large part, or even all, of human creativity can be understood as a novel combination of pre-existing ideas or objects *(Boden, 2010)*.

Latent space is a good example for demonstrating computational creativity. Not only because of its combinatorial power, but also because of its mathematical possibilities of **interpolation** and **extrapolation**. To decipher this, let's take a look at the concept of latent space.

Interpolation refers to the calculation of values between two data points, extrapolation to the calculation of values that can be derived from the pattern of data points *(Royal Academy of Arts, 2018)*.

2.4.1 The computational creativity of the latent space

Latent space can be defined as a multidimensional **vector space** that represents the learned features of a data set. Multidimensional vector means that
the latent space can encode different feature dimensions such as age (vector old - young), gender (vector male - female) or facial expression (vector sad - happy) in a data set of faces. It serves as a conceptual framework that facilitates the representation and processing of

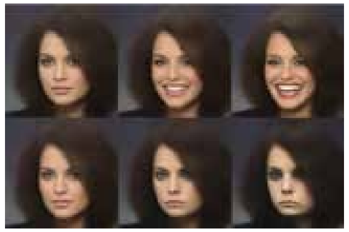


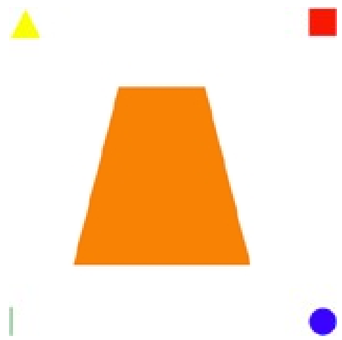
Fig. 6: The original image (left) is interpolated in Latent Space in the direction of a "smile vector", creating a series of versions of the original, from smiling to sad.

complex data and facilitates effective analysis and modeling. The representation of the latent space is derived from the input data and is characterized by
The extraction of meaningful features, which improves the ability of AI systems to understand and process information. This concept is crucial for AI as it enables the efficient representation and processing of complex data.

In GANs, VAEs and diffusion models, latent space is used for image generation. By manipulating points in latent space, diverse and coherent images can be generated, demonstrating the ability of latent space to encode a wide range of variations *(Kingma & Welling, 2013)*. In language models, latent coding of textual data facilitates information retrieval by enabling the identification of semantically related documents or passages. The theoretical understanding of latent space is still evolving. While there are no definitive theories, various perspectives and assumptions help to understand its role. For example, the Manifold Hypothesis states that low-dimensional latent space is often responsible for the high-dimensional observations that
we see. This means that highly complex phenomena of reality can be reduced to a few **virtual** patterns. Although this concept is only roughly defined, it works in practice *(Kingma & Welling, 2013)*.

We can navigate this space to generate new combinations of features that lead to new images or designs. Exploring latent space allows for both aleatory (random) and structured creative working, providing a new framework for understanding creativity.

The phenomenon of **AI** hallucinations, which are not hallucinations in the traditional sense, refers to the unique and often unexpected results produced by AI systems when they interpret and recreate data in ways that humans cannot predict. This phenomenon was first
observed in language models and computer vision tools and describes the recognition of patterns or objects that are non-existent or imperceptible to human observers, producing nonsensical or completely inaccurate results *(IBM, n.d.)*. While AI hallucinations are certainly *undesirable* outcomes in most cases, they also offer a number of fascinating use cases that can help artists explore the creative potential of AI.



Instead of describing an orange trapezoid in detail, we find it as an interpolation of the low-complexity elements of the two dimensions shape and color in the latent space of the data points green line, yellow triangle, red square and blue circle.

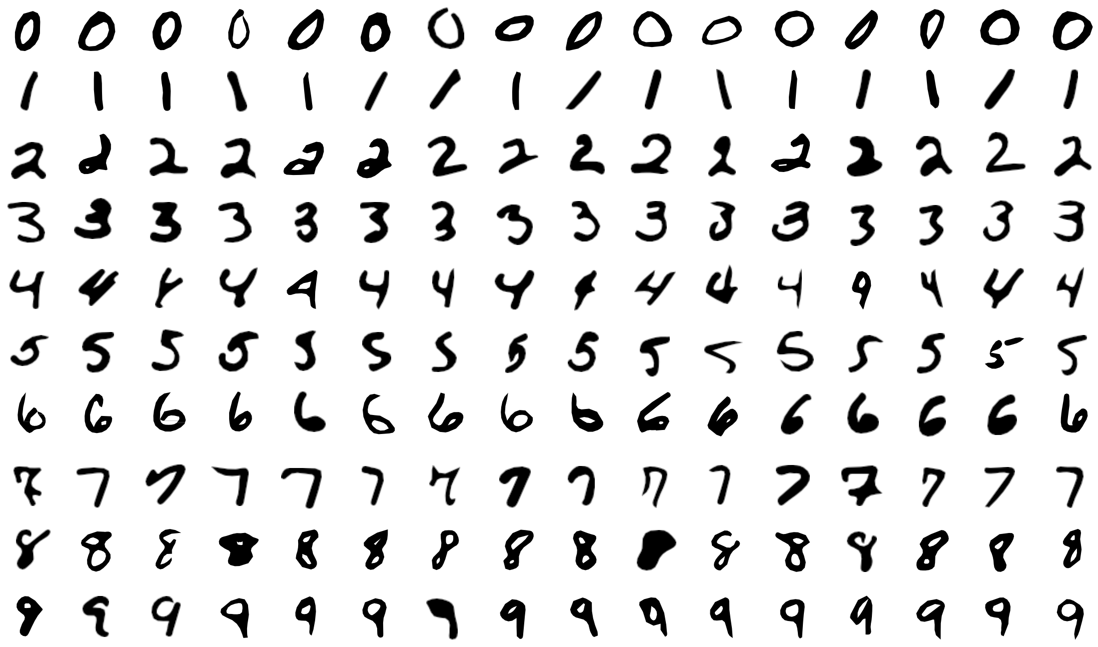


Fig. 7: Excerpts from the original MNIST *(Modified National Institute of Standards and Technology)* dataset. This is a database of handwritten digits with a training set of 60,000 examples.

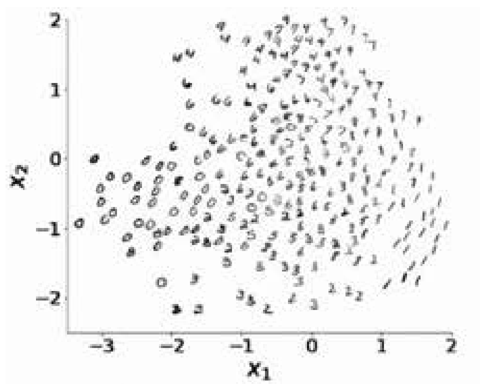


Fig. 8: After training, the digits can be grouped using latent variables.

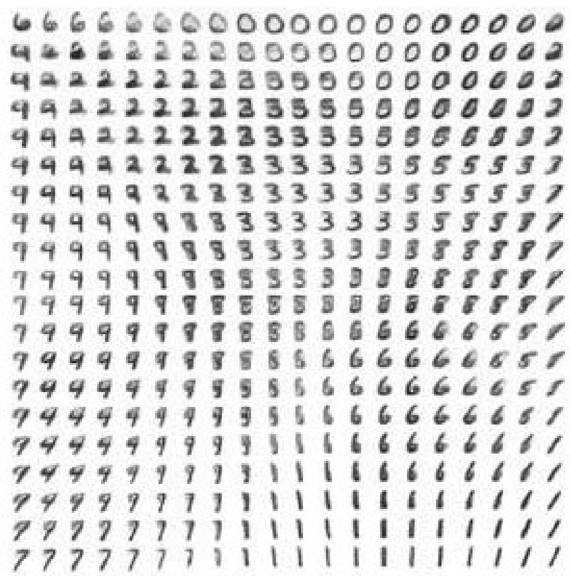


Fig. 9: Latent space of the MNIST dataset.

potential of AI in a positive way, e.g. by discovering new ways or perspectives *(The AI GRID, 2024)*.

2.4.2 Metacreativity as a hermeneutic of the AI era

The term **meta-creativity**, coined by Eduardo Navas, refers to a form of creativity that goes beyond human boundaries by including non-human systems in the creative process. In this view, machines not only participate in creativity, but also expand the scope of what is possible.
Navas argues that AI-generated art represents a cultural shift that blurs the boundaries between human and machine creativity and challenges traditional notions of authorship and artistic genius *(Navas, 2022)*. Navas defines meta-creativity as

> *"a cultural variable that emerges when the creative process moves beyond human production to include non-human systems. This definition includes artificial intelligence and machine learning; for emerging intelligent technology, specifically, this means that a non-human entity is able to 'learn' in order to produce something that appears creative" (Navas, 2022, p. 231).*

Similar to the idea of computational creativity, meta-creativity does not explicitly state that something is creative, but that it appears creative. The challenge is therefore obviously how the results of creative machines are evaluated and interpreted by humans, since they are produced by non-human actors.

In general, the value of a painting seems to decrease as soon as humans learn that it was created by AI *(Horton, White & Iyengar, 2023)*. The sale of the AI-generated portrait "Edmond de Belamy" for $432,000 at the 2018 Christie's art auction was the exception rather than the norm, but has raised questions about how credit and responsibility should be assigned to the people involved and how the anthropomorphic perception of the AI system (the algorithm was credited as the creator) contributed to the success of the artwork *(Epstein, Levine, Rand & Rahwan, 2020)*.

The notion of creative works by machines brings a new level of complexity to the debate on creativity, as it raises the question of what creativity actually is and what counts as creative.

Fig. 10: The portrait "Edmond de Belamy", the algo- rithm is named as the author in the bottom right corner.

3 Conceptualize creativity

Creativity, which has long been considered a hallmark of human intelligence, has been defined as the ability to develop ideas or products that are both novel and valuable in a particular cultural or social context *(Runco & Jaeger, 2012)*. Traditionally, creativity is seen as a deeply personal and emotional process that requires intuition, empathy and cultural sensitivity.
This process often reflects the artists' individual experiences, which are influenced by emotions, social conditions and the desire to express something profound or original *(Runco, 2023)*.

In the 18th century, Immanuel Kant, in his work on aesthetics, considered creativity as a form of self-expression inherent to **the genius** of the individual, thus laying the foundation for the modern understanding of the artist as an autonomously creative person *(Kant, 1790)*.
This focus on the individual's ability to go beyond con- ventions and create something new and meaningful laid the foundation for the modernist approach to creativity.

3.1 Creativity in the modern age, postmodernism and metamodernism

At the beginning of the 20th century, **modernism** emerged as the predominant artistic movement, characterized by a belief in progress, innovation and overcoming traditional boundaries. Modernist creatives, such as Pablo Picasso and Virginia Woolf, strove for completely new forms of expression, emphasizing originality and a departure from the conventions of the past. Creativity in modernism was associated with the genius of the artist, who was seen as a visionary capable of revealing truths through novel techniques and ideas. According to Buehrle *(2008)*, art was clearly no longer a reflection of the world, but reflected the subjective view of individuals and creatives.

Postmodernism emerged in the middle of the 20th century as a reaction to the ideals of modernity. Postmodern theorists rejected the idea of abso-
The post-modernist movement rejected the "unambiguous truths" and grand narratives and emphasized fragmentation, plurality and irony in creative work *(Aylesworth, 2015)*. In postmodern thinking, the artist-creator became a "bricoleur*" (cf. Lévi-Strauss, 1962, p. 16)*, constructing meaning from pre-existing cultural material rather than developing their own ideas *(Duy- medjian & Rüling, 2010)*. This shift was particularly exemplified in remix culture, where the boundaries between original creation and appropriation were blurred and the modernist ideal of the a r t i s t as lone genius was challenged *(Navas, 2012)*.

At the beginning of the 21st century**, metamodernism** is a framework that alternates between the ideals of modernism and postmodernism. Artistic works and theories of metamodernism attempt to reconcile the sincerity and idealism of modernism with the irony of postmodernism. As Vermeulen and Van den Akker *(2010)* describe, metamodernism is characterized by a hopeful search for meaning, even as it acknowledges the impossibility of ever fully achieving that meaning.

In metamodernism, creativity is seen as a balancing act between serious exploration and cynical self-reflection, in which artists are aware of the cultural and social constructs that shape their work and yet strive for an authentic expression. Moreover, in the post-humanist discourse, the boundaries between man and machine are becoming increasingly blurred, and it is not uncommon for machines to be used as well.

Fig. 11 The Fountain by Marcel Duchamp is considered a key work of modernism.

Fig. 12: Keith Haring combined mass culture with high culture and thus became a postmodern icon.

Fig. 13: The Elphilharmonie is a meta-modern landmark that combines irony and tradition.

creativity - or at least the potential for it *(Berlich, 2022)*.

When we integrate artificial intelligence and computational creativity into this theoretical framework, it becomes clear how they challenge traditional, anthropocentric notions of creativity. While the concept of artistic genius was central to modernism, it is increasingly being challenged by artificially generated works. Postmodernism already challenged the idea of originality, but AI complicates this discourse by creating works that are not mere copies, but innovative combinations of existing data in latent space with almost no human intervention.

Metamodernism, with its focus on the fluid transition between irony and seriousness, sense and nonsense, as well as the deconstruction of authorship, offers a suitable framework for analyzing artificial creativity. The use of AI in art alternates between the cynical view that it is merely an algorithmic imitation of human creativity and the serious conviction that AI not only imitates the creative possibilities of humans, but is creative in its own right. The collaborative nature of AI-generated art, where the human curates and directs, fits into the ethos of metamodernity, which embraces both human agency and technological potential *(Berlich, 2022)*.

Fig. 14: Not only metamodern, but also meta-creative: Boris Eldagsen's "The Electrician".

3.2 Human vs. machine creativity: Critical comments

So how "creative" can machines be? Margaret Boden *(2014)* argues that the answer is not technical, but philosophical. Boden argues that while AI systems can produce results that appear creative, they do not have the intentionality, emotionality or conscious understanding required for true creativity.

Critical voices such as David Gelernter *(1994)* are of the opinion that AI-generated art is technically impressive, but does not have the emotional resonance and cultural significance of works created by humans. Gelernter emphasizes that art is not just about the end product, but also about the process and the emotional journey of the artist. Machines that lack emotion would not be able to replicate this essential component of the creative process *(Gelernter,*

1994, S. 149). Such **demarcation mechanisms** are also frequently heard in the art scene itself: Artists such as Aelfleda Clackson argue: "Unlike AI, I have feelings, and that is what makes my work meaningful." *(Clackson, 2024)*.
Catherin Botha *(2023)* accuses Boden - and thus the criticism that refers to the distinction between man and machine - of **anthropocentrism** with regard to her definition of "genuine" creativity and encourages us to disregard human uniqueness:

> *"we need not jealously guard creativity as an ability that is exclusively held by humans. If we recognize the anthro-pocentric conception of agenthood as problematic within our understandings of creativity, it seems possible to grant non-human agents and artefacts entry into the world of creativity, without undermining the value of the artefacts and ideas created, and without undermining the worth and dignity of our humanness. Just as it is easy for us to say a system is logical, or even rational in its activities, and so it seems to me that saying a system is creative, should be just as easy" (Botha, 2023 p. 209).*

3.3 The Latent Space as a meta-creative medium

In this context, the term "medium" is understood as an element that makes the world "out there" accessible to us through **abstractions**. This definition according to Vilém Flusser *(2000)* considers images in particular to be this medium. Flusser distinguishes between traditional and technical images. Although latent space is a mathematical abstraction, the concrete images that arise from it are of a technical nature: they are made accessible to us through a "Apparatus" - the algorithm - in a format that we can understand. In his book "Illustration - 100 Ways to Paint a Bird", Felix Scheinberger *(2013)* dedicates an entire chapter to computers and generative design and clearly assigns images created with an algorithm to illustration. Even if AI can generate images that are astonishingly similar to photographs, it is in their nature that they are not. As a meta-medium, AI can simulate traditional techniques, but generated images are ultimately images created with a program and therefore illustrations.

So how can a medium be creative? After training an AI, there is a file that contains the training material in is codified in a latent mathematical space. While the human mind is continuously updated *(see Neuroplasticity chapter 2.1.1)*, the latent space after training is a crystalline space of fixed values. However, it also contains all possible combinations of these v a l u e s , and therefore also new and less new, or creative and non-creative solutions. The latent space can therefore be seen as a representation of all creative possibilities in this data set. This makes the latent space meta-creative in two senses: firstly because it is the result of non-human learning and secondly because it already **contains** all the solutions **and does not discover them**. It is the human who *remains* basic creative in this sense, because he has to search for and uncover the solutions in the latent space, but not because he generates them.

In the stable diffusion community, the term "skill issue" describes people who fail to generate the desired image from the latent space as being insufficiently familiar with the program. The failure is attributed to a lack of skills, not technical limitations.

3.4 Creativity as a process of curation and contextualization

When working with image generators, the role of the creatives shifts from direct creation to curating the machine-generated results by selecting and refining the AI-generated works that correspond to their artistic vision. This process mirrors the **oscillation** *(cf. Vermeulen, Van den Akker, 2010, p. 1)* between control and surrender, where creatives navigate the delicate balance between human intention and the futility of it.

The act of curating introduces a new dynamic into the creative process, where the artist-creator acts as an intermediary between the algorithmic possibilities of AI and the cultural meaning inherent in the final work. This dynamic reflects the general shift in creative practice, in which the role of the creator is no longer limited to the production of new works, but also the selection, refinement and reinterpretation of machine-generated content. It is the task of artists to select and contextualize the creatively relevant ideas of the machine *(Navas, 2021)*.

4 The portrait in Latent Space

"The first painting was a portrait," writes Joanna Woodall in the foreword to her book, "Portraiture: Facing the Subject", referring to the story of Narcissus, a young Man falling in love with his own reflection *(Woodall, 1997, p. 1)*. Portraiture has always played an important role in art, serving as a medium for exploring human identity, emotions and social status. As Shearer West describes in her book "Portraiture", portraits have long been used to immortalize the individual, such as the early form of portraiture, the totem masks in ancient Rome. By depicting not only the physical likeness but also aspects of the sitter's psychological and moral character, portraits became not only a means of ancestor worship but also a means of representation *(West, 2004)*.

From the Renaissance onwards, portraiture developed to capture the essence of the sitter's inner world, reflecting the growing interest in individualism and self-knowledge *(Woodall, 1997)*. Artists such as Leonardo da Vinci and Rembrandt mastered the difficult balance between realism and idealization, making the portrait both a representation of the person and a broader statement about humanity and society.

The self-portrait was also a way for creative people to explore their own identity and their place in the world. The self-portrait became a method of self-reflection and artistic experimentation, as can be seen in the works of Albrecht Dürer, Vincent van Gogh and Frida Kahlo. These artists used their own image to express personal struggles, aspirations and existential questions, contributing to the cultural significance of the portrait as more than just a physical likeness, but a psychological and emotional exploration *(Pointon, 1993)*.

Fig. 15: Caravaggio: Narcissus.

The portrait also offers the advantage that the viewer can position themselves emotionally in relation to it. Unlike nature photographs, for example, a portrait is accompanied by an emotional reaction. This can even lead to identification with the figure, especially in the case of average and simple portraits. *(McCloud, 2001)*.

In the context of AI-generated art, the portrait offers a fascinating juxtaposition between the human uniqueness traditionally captured in thousands of portraits (which are also part of the training dataset) and the artificial creativity of machines. Both the emotional depth and the symbolic significance of the portrait genre make it a compelling Case study for exploring the possibilities and limits of AI-driven creativity.

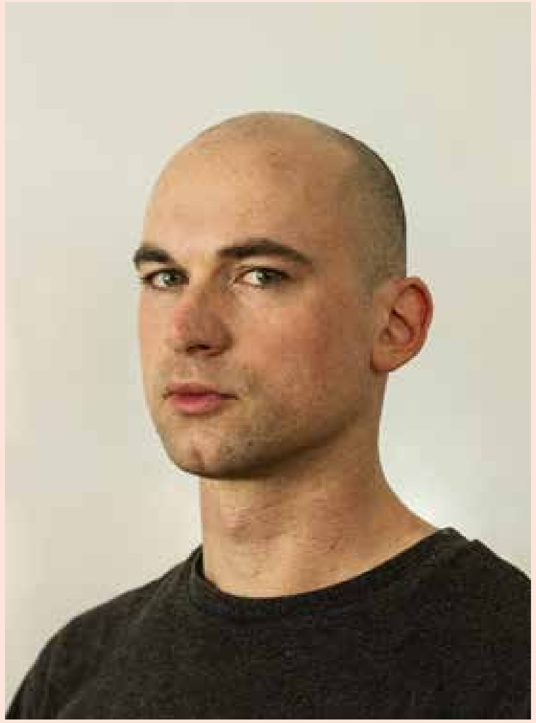
Fig. 16: The contemporary artist Bryan Lewis Saunders, for example, uses self-portraits to document the effects of various drugs on his body, his perception and ultimately his creative work.

4.1 Automation of creativity: Generative design techniques

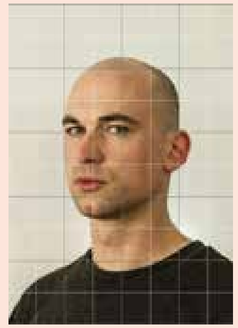
This work is exploratory and therefore makes no claim to completeness. In reference to the creative space a s described by Thomas and Martin Poschauko in their work "Nea Machina" *(2013)*, I have created a (self-)portrait and a prompt (in German "Eingabeaufforderung", the English name for the text description of the image to be generated), both of which serve as **formal constants** of the creative experiment. The self-portrait for this comes from a photo series that I created at the beginning of the 2020 pandemic. These two elements can vary in terms of content, although the generative self-portrait, and thus the approximate pose and approximate content, remain the same. The Experiment is thus based on a series of portraits that are structured according to certain guidelines such as design techniques and design principles. Similar to Butler, Holden, & Lidwell *(2010)*, I have collected different principles that occur again and again in creative works or have played a role. I use this list to explore the possible applications of these creativity techniques in Latent Space. This list is not exhaustive, but only serves as a rough guide for new perspectives and creative experiments. At the same time, I allow myself to experiment outside the series in order to come up with additional ideas once I have exhausted the content of an idea for the time being.

For my experiments, I use the Stable Diffusion program described at the beginning for image generation and ChatGPT for textual work. Prompts and further settings are described for each experiment. The models of the AI programs have repeatedly changed or been updated during the course of the experiment - this also results in a changed latent space, as the models can differ in terms of their training data. However, even if the experiments differ minimally in terms of style, the techniques used remain the same.

This approach is also described by Poschauko and Poschauko as a creative space with a "back door" *(Poschauko & Poschauko, 2013, p. 13)*



The **portrait** is upright, as are most of the experiments. As some models have different technical requirements, different resolutions of 512x512 pixels, 1024x768 pixels and 1024x1024 pixels are prepared as possible image prompts (see page 42). For some experiments Separate templates, such as grids or cut-outs, are created for the individual elements.



This part can usually be omitted, as most models have been trained on predominantly photographic material and the image aesthetics therefore correspond more closely to a photograph.

The **prompt** is the image description with which I can create an image similar to the one above. This prompt consists of a medium, a subject, an item of clothing and a description of the surroundings. For some experiments, only the subject is used to minimize stylistic influences.

A photo of
a young bald man,
black t-shirt,
white background

Additional information is optional, and is used depending on how closely I want to stick to the initial concept.

4.1.1 Technical settings of diffusion models

The following descriptions of attitudes and control options are based on Andrew Wong's *(2024a)* description of the diffusion process; other sources are mentioned in individual cases.

Diffusion steps

Stable Diffusion is basically a text-image model that generates images based on a prompt. It belongs to a class of AI models known as diffusion models. These are trained on large datasets where each image gradually fades into (pixel) noise through a process known as forward diffusion. The model learns to *reverse* this process *(reverse diffusion)* in order to reconstruct the original image. If the model masters this reversal, it can generate new, coherent images from random noise that resemble the training data.

To create an image, Stable Diffusion starts with a completely random pattern in a latent space. The model then uses the process of reverse diffusion to gradually remove the noise from this initial pattern so that an image gradually emerges, which corresponds to the given prompt. Since Stable Diffusion is a **conditional** image generator, it uses the prompt to control this denoising process. In this way, the model can generate different images from the same output noise depending on the prompt.

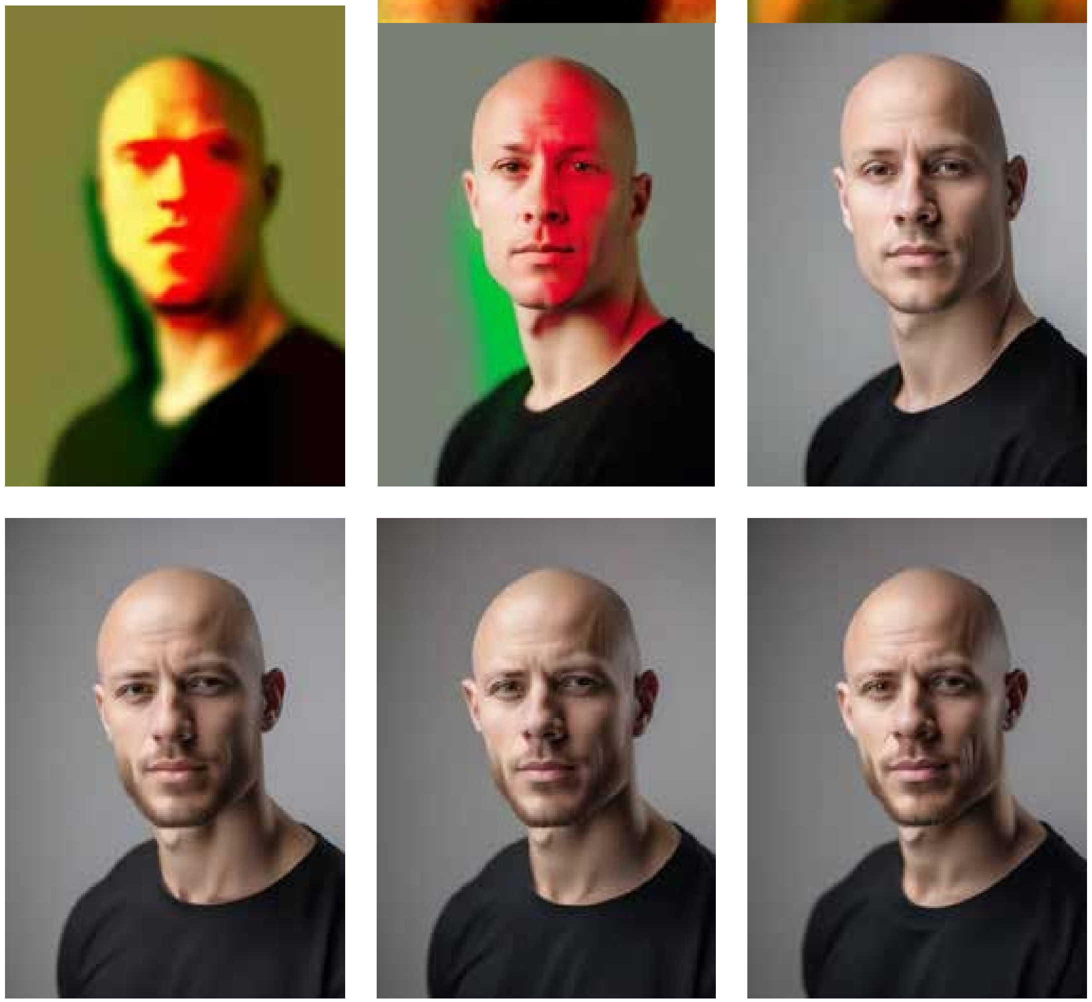
The Stable Diffusion generation process includes an algorithm that controls how the noise is removed at each step. At the beginning, larger steps are taken to determine the global structure, such as composition or color distribution of the image. As the process progresses, the steps become smaller in order to refine smaller details such as textures and edges. This gradual refinement means that the image begins to stand out from the noise in the first steps, but remains blurred. With each step, the image becomes clearer, and from about 20 steps onwards, the model concentrates only on improving the fine details instead of changing the overall composition.

By adjusting parameters such as diffusion steps (which determine the number of iterations to refine the image

Unconditional image generators produce a random product. The first image generators, such as the GANs or VAEs, were therefore more difficult to control, as the points in the latent space had to be laboriously found.

Many mathematical methods of the denoising process lead to the model converging after about 20 steps. This means that
stabilize the results and approximate a certain state or distribution that represents the underlying data well. There are also methods, the so-called samplers (see page 36), which converge after just a few steps. The Latent Consistency Model Sampler can converge images after just one step.

`Prompt: A photo of a young bald man, black t-shirt, white back-ground Steps: (0-33), Sampler: Euler a, CFG scale: 8, Seed: 3319583382, Size: 768x102`



The portraits produced range from detailed, almost photographic-looking results to highly impressionistic, blurred compositions.

Seed

The random noise pattern that is gradually denoised in Stable Diffusion is first initialized with a number called a seed. This seed is not completely arbitrary; it can be set manually or generated randomly by the system. Using a random seed is advantageous if I want to examine a wide range of images generated from the same prompt, as this leads to some variability. However, if I specify a particular seed, I can generate reproducible images. This feature is particularly useful when experimenting with other parameters, making slight changes to a prompt or sharing my work with others who can reproduce the exact same image. The main advantage of using a fixed seed is that image generation with the same prompt, the same parameters and the same seed will produce exactly the same images every time.

This consistency allows me to create multiple similar variations of an image, perform controlled experiments or ensure that certain results can be reliably reproduced.

In the stable diffusion community, some seeds have been identified that have a higher probability of generating images with certain characteristics, e.g. certain color palettes or compositions. Knowledge of these seeds enables users to achieve more targeted results, which can be helpful when searching for specific artistic results or imitating a particular style *(Inglewood, 2023)*.

Top right:

`Prompt: A photo of a young bald man, black t-shirt, white background Steps: 20, Sampler: UniPC, CFG scale: 8, Seed: 849238882843, Size: 768x1024`

Bottom left:

`Prompt: A photo of a young bald man, black t-shirt, white background Steps: 20, Sampler: UniPC, CFG scale: 8, Seed: 100, Size: 768x1024`

Bottom right:

`Prompt: A photo of a young bald man, black t-shirt, white background Steps: 20, Sampler: UniPC, CFG scale: 8, Seed: 1, Size: 768x1024`

CFG scale

In simple terms, the CFG scale *(Classifier Free Guidance scale)* or guidance scale is a parameter that controls how closely the image generation process follows the prompt. The CFG-Scale essentially strikes a balance between output accuracy and creative freedom, allowing users to specify exactly how literally the AI interprets the input text. At higher CFG values, the model adheres more strictly to the prompt and produces images that more closely match the specified details.
However, at very high CFG values, the images can be taken too "literally", which can lead to potential problems such as oversaturation, loss of detail or artifacts, especially if the prompt is too restrictive or does not contain enough descriptive detail. Finding the optimal CFG value is often a matter of experimentation and balancing prompt accuracy and image quality. Typically, a CFG value between 7 and 11 is considered ideal for most applications as it strikes a balance between maintaining prompt accuracy and maintaining high visual quality. If the CFG value is set too high the generated image may suffer from excessive sharpness or noise; if it is set too low, the output may drift too far from the intended prompt.

In the experiment conducted with Stable Diffusion, the portraits were generated by adjusting the CFG scale to control the influence of the prompt. With a higher scale, the AI produced highly textured, photorealistic portraits, while a lower scale resulted in flat, desaturated images. A CFG scale that was too high resulted in garish, oversaturated images.

`Prompt: A photo of a young bald man, black t-shirt, white back-ground Steps: 20, Sampler: UniPC, CFG scale: {1-30}, Seed: 100, Size: 768x1024`







It is also possible to hack the program. With knowledge of Python, the source code of Stable Diffusion can be edited and negative or no CFG values are possible. The generated images i g n o r e the prompt and random images can be created. Without the CFG scale, Stable Diffusion behaves like an **unconditional** image generator (see page 28).

Weight (Weight)

In the context of stable diffusion and other machine learning models, word *embeddings* are numerical representations of concepts - such as words, sentences or images - that capture their meanings in latent space. How strongly the model emphasizes attention to certain embeddings in the prompt can be changed with the help of word weightings.

For example, with a prompt such as `(sunset:1.5)(beach:0.75)`, the model prioritizes "sunset" (weighting 1.5 times) over "beach" (weighting 0.75 times), resulting in an image with more sunset elements, but still some beach aspects. However, the attention does not necessarily mean that an element becomes larger; it can also simply become more prominent or more detailed.

There are various techniques for using weightings *(catboxanon, 2024)*:

- Numerical weights are given in the prompt to fine-tune attention to specific words. `(dog:2)` doubles the focus of the model to "dog", while `(cat:0.5)` shifts the focus to "cat" around the half.
- With multiple concepts such as `((man:0.5), (woman:0.9), (frog:1.2))` the concepts in the image weighted. `You` are less in the foreground, but `woman` is more and `frog` gets the most attention.
- Concept blending: `(man:1.5|woman:1.5|frog:1.5)`, blends the concepts together during image creation. This technique can be used to create images with properties of all concepts.
- Concept change and mixture: With `[man:frog:0.25]`, the image will change from an image of a man to an image of a frog after the first quarter of the diffusion steps. This means that the rough composition (early steps) is based on `man`, but the details (later steps) are based on `frog`. This method is suitable for creating hybrid images.

Negative word weightings can also be used to reduce the influence of a word. Too much weighting, on the other hand, leads to oversaturated images or distorted features.

An embedding consists of a tensor, i.e. a mathematical object in latent space. This tensor consists of 768 numerical values that correspond to individual dimensions; these are not spatial dimensions such as height, length and depth, but mathematical dimensions. Embedding is therefore the mathematical abstraction of a (linguistic) concept. The embedding `frog`, for example, is codified using this tensor:

`tensor([-0.0040, 0.0116, -0.0175, ..., 0.0014, 0.0149, 0.0066]).`

When weighting, this tensor can simply be multiplied or offset against other tensors.

The syntax of special characters such as `()`, `[]` or `:` varies depending on the user interface. This is how Stable Diffusion `(frog:1.5)` `frog++` uses them, `frog::1.5.` Cogniwerk, Midjourney

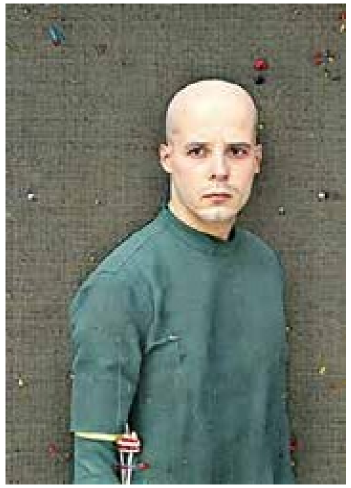
A standardized spelling has been has not yet been implemented.

`Prompt: A photo of a young bald (man:{-8-8}), black t-shirt, white background, Steps: 20, Sampler: UniPC, CFG scale:8, Seed: 100, Size: 768x1024`

If an embedding is given too much negative or positive weight, this can lead to distortions and artifacts that can be used artistically.

Top left: `(man:-6.0)`

Top right: `(man:6.5)`



A concept mix usually only works as well as the data set allows. A mixture of `man`, `woman` and `frog` leads to a group foto, as the concepts are more associated with *individual entities*. In other words, for example, `man`, `makeup` and `beard` produce different results,

because these embeddings with *properties* be associated.

`Prompt: A photo of a young soon ((man:0.5), (woman:0.9), (frog:1.2)), black t-shirt, white background, Steps: 20, Sampler: UniPC, CFG scale: 8, Seed: 100, Size: 768x1024`

A mixture of `man`, `woman` and `frog` on the other hand, leads to a female read person with slight wrinkles, secret corners, a larger nose and large eyes, as well as a frog T-shirt. How the prompt is interpreted depends on the seed and can therefore vary from seed to seed.

`Prompt: A photo of a young bald (man:1.5|woman:1.5|frog:1.5), black t-shirt, white background, Steps: 20, Sampler: UniPC, CFG scale: 8, Seed: 100, Size: 768x1024`

When the concept is changed, `man` is generated for five steps, but then switches to the `frog` concept for the remaining 15 steps. This then directs the details in the direction of `frog`.

`A photo of a young bald [man:frog:0.25], black t-shirt, white background, Steps: 20, Sam-pler: UniPC, CFG scale:8, Seed: 100, Size: 768x1024`

Sampler

Samplers play a crucial role in image generation as they determine how pixel noise is removed from the initial random input to produce a clear and coherent image. As samplers are **mathematical methods** for noise removal, several approaches have been developed in parallel in recent years, affecting both the speed of generation and the quality of the final product. The samplers differ in their approach and performance *(Wong, 2024b)*:

- **Euler** and **Heun** are classic methods for solving differential equations, whereby Euler is fast but less detailed and Heun offers more accuracy at a lower speed.
- **LMS** and **PLMS** improve accuracy through averaging of steps, but can cause noise at lower steps.
- The **DPM family**, including DPM, DPM2 and DPM++, reduces the number of steps required for high quality output, with DPM++ providing an efficient balance between speed and accuracy.
- **DDIM** delivers high-quality images quickly, but requires more steps for optimum results.
- Ancestral samplers like **Euler-A** introduce random noise into the process and produce creative but less stable results.
- The newer UniPC **sampler** combines an efficient Prediction and correction to achieve high-quality results in just a few steps.
- **LCM** is also a new sampler that can be used with just a few steps and can generate an image after just 1-4 steps. This is often used for real-time applications.

Choosing the right sampler depends on the desired balance between speed and image quality.
but can also be chosen as a conscious decision for a certain aesthetic.

Euler and Heun are proper names, the other names are abbreviations.

LMS: Linear Multistep
PLMS: Pseudo-Linear Multistep
DPM: Denoising Probabilistic Model
DDIM: Denoising Diffusion Implicit Model
UniPC: Unified Predictor-Corrector
LCM: Latent Consistency Model

`Prompt: A photo of a young bald man, black t-shirt, white back-ground, Steps: 20, sampler: {selection}, CFG scale: 8, Seed: 4829334432, Size: 768x1024`

As the LCM sampler requires fewer steps and a smaller CFG scale, the values for this prompt are different:
`Steps: 4, CFG scale: 1.5`

PLMS

DPM

DDIM

Euler A

UniPC

LCM

Denoising strength (Denoising Strength)

The denoising strength can be used to influence the strength of the noise removal during image generation. This parameter ranges from 0 to 1: A value of 0 retains the original noise, while a value of 1 completely converts the input and generates a completely new image.
Instead of starting with random noise, an image can also be used as a prompt. With a denoising strength of 0, the input image remains intact, while a new image is generated with a value of 1. With between values, mixtures of random and predetermined values can be achieved *(Wong, 2024c)*. Several application modes have developed from this: Image-to-image (see also page 42) uses this technical fact directly to influence the color distribution as well as the composition by means of image prompts. Thanks to special samplers such as the LCM, which make it possible to create a detailed image, applications have been created in which the user "paints" in the input layer and the result is visible in real time in the output layer.

With inpainting, only the areas of a picture in which the diffusion process is to take place are marked. This allows selective changes to be made to a picture.

`Prompt: A photo of a young bald man, black t-shirt, white back-ground Steps: 20, Sampler: UniPC, CFG scale: 8, Seed: 100, Size: 768x1024`

`Denoising Strength: {0.0-1.0}`

The initial original noise is always "gray" as it is a normal distribution of RGB pixels. However, this also means that images can never be completely black or completely white, as Stable Diffusion does not allow large deviations from this normal distribution. With a hack, the so-called noise offset, the original noise can be shifted towards black or white, which allows for much higher-contrast images.

4.1.2 Exploring the latent space

In the context of portraits, certain aspects of creativity can be automated, especially the combinations of features. The simplest form of automation is to run through a set of data. Since we can navigate through the latent space using text input, concepts or properties such as material, hair color or similar can be defined linguistically in a **variable** and different versions generated: Prompt: `A portrait of a person made out of {mate- rial}`.

The collection variable `{material}` can then contain all possible materials such as leather, glass, metal or cardboard. When using several variables
This concept can be extended with the help of simple combinatorics. However, there are other ways of exploring latent space in a controlled manner. The most common ones are presented below.

Text prompt
A prompt is a way of communicating with a machine - and in the case of image generation, essentially a form of **image description**. With the integration of CLIP, the use of text prompts became the most common method for controlling image-generating AI models.

CLIP *(Contrastive Language-Image Pretraining) is* a model developed by OpenAI that learns visual concepts from natural language. It works by pairing images with text and training the model to associate descriptions with the correct images. CLIP enables tasks such as image generation, classification and image description by understanding language and images together *(Radford et al., 2021)*. Since the majority of the data material used for training consists of English image-text pairs, it makes sense to prompt in English in order to obtain predictable results.

This practice, known as "prompt engineering", quickly developed into a sophisticated craft from 2022 onwards. A typical prompt is a detailed and precise description of an image and consists of several embeddings.

There are also some non-English words that CLIP can also assign. However, these are usually words that are frequently used internationally (i.e . on the Internet), such as "master" and "kebab", "Aam Aadmi" (Indian equivalent for "Max Musterman", also the name of a political party in India).

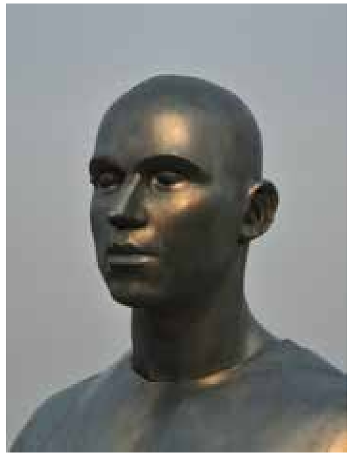
Each embedding has a weight of 1, but the earlier it appears in the prompt, the more influence it has.

Therefore, the prompts often follow a general syntax such as "describing the subject or action, giving additional details, in a certain style".

Different training methods have led to variations in the creation of prompts, ranging from descriptions in natural language to keyword-based approaches. A common technique is to use certain keywords to achieve a desired style. For example, terms such as "hires", "8K" and "sharpness" enhance image quality, while mentioning camera models can evoke certain aesthetic qualities associated with them.
However, it is not possible to reproduce exact visual properties: While a prompt can define the subject, it has limited control over composition and the reproducibility of finer details.

In addition, "word classes" have developed so that a prompt can be structured in a more targeted way: With a syntax such as `{format} of {subject term}, a {so- lidifier}, {perspective}, by {artist:weight} | by {artist:weight} / in the style of {style}, {va- riant}, {booster}, {exclusion}`, a prompt can be specified in more detail. However, these word classes have no equivalent in the model, but merely serve as an orientation for describing an optimal prompt. With the advent of Large Language Models *(LLMs)* such as ChatGPT, however, it has become increasingly common to define the rough structure and let the LLM generate the optimal prompt.

`Prompt: A statue of a young bald man made of Polycarbonate, Steps: 20, Sampler: DPM++ SDE Karras, CFG scale: 8, Seed: 2757207243, Size: 768x1024`



`Prompt: A statue of a young bald man made of Metal, Steps: 20, Sam- pler: DPM++ SDE Karras, CFG scale: 8, Seed: 2757207243, Size: 768x1024`

`Prompt: A statue of a young bald man made of Plywood, Steps: 20, Sampler: DPM++ SDE Karras, CFG scale: 8, Seed: 2757207243, Size: 768x1024`

A Con- trolNet (see page 43) was also used for this series in order to to stabilize the position of the head and merely replace the material.

Image prompt

An image prompt directs an AI model to transform an existing image or generate variations, combining both visual input and text. Unlike text-only prompts, where the descriptions only control the creation, image prompts provide a basis for color and composition, with the model adjusting or enhancing the image according to the text. As the base image provides the structure and composition, the text prompt can instruct changes in style, detail or content.
Terms such as "watercolor" or "mixed-media collage" change the visual style, while the denoising strength controls the degree of change and enables either subtle adjustments or dramatic changes. An image prompt can therefore be used as a visual anchor. The procedure for inpainting is similar, except that, as mentioned above, it is limited to a previously defined area.

`Prompt: A photo of a young bald frog, black t-shirt, white background, Steps: 20, Sampler: DPM++ 2M Karras, CFG scale: 8, Seed: 100, Size: 768x1024`

`Denoising Strength: 0.7`

ControlNet

ControlNet is a method used in AI image generation to gain additional control over certain aspects of the output by adding structural information, such as
z. e.g. pose, depth or line drawings, can be included in the generation process. In contrast to regular image or text-based prompts, ControlNet enables the model,
follow detailed guidelines when generating images *(e.g. hand-drawn sketches, human poses or 3D data)*, ensuring more precise control over the composition and structure of the output.

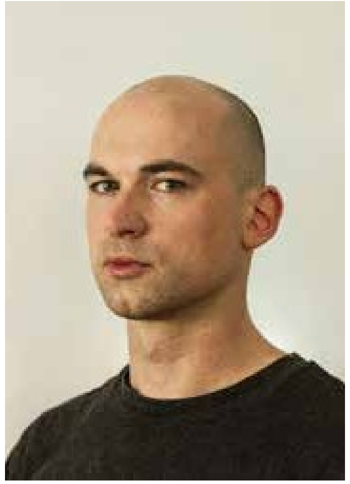
In ControlNet, the visual input *(e.g. a line drawing or a body pose)* acts as a blueprint, ensuring that the generated image matches this template exactly. The prompt, on the other hand, adds style, content and further *(content-related)* details and enables both creative and structural adjustments. For example, I can specify a pose and use `a` prompt such as `a futuristic robot in a forest` to ensure that the generated figure adopts the exact pose while adding the desired content elements.

This technique allows far greater control over the outcome and combines the precision of visual guidance with the creativity of prompts. ControlNet's ability to maintain both structure and style makes it ideal for tasks where a controlled or precise outcome is critical *(Wong, 2024d)*

As a submodule, ControlNet has its own models (and its own settings!), which themselves also have their own latent space. This means that the way in which an input is interpreted also differs between different ControlNets. The selection of current ControlNets and their description is beyond the scope of this paper,
as there are already several hundred types and use cases of ControlNets. The most common are line drawings (SoftEdge, Canny, Scribble or Sketch, MLSD, Lineart) as image support, body poses (AnimalPose, OpenPose), depth or 3D data (Depth, Normal) or reference data (Reference, Revision, IP-Adapter, Instant-ID, T2I-Adapter, Photomaker).
There are also special cases that are more like filters, such as Recolor or Blur, which either add color to a monochrome image or soften it (the latter is a form of image prompt in which colors and shapes are roughly adopted).

`Prompt: A futuristic robot in a forest Steps: 20, Sampler: DPM++ 2M Karras, CFG scale: 8, Seed: 100, Size: 768x1024`

`ControlNet Model: sdxl-scribble-Anyline, Weight: 1, Threshold A: 0.5, Threshold B: 0.5, Guidance Start: 0, Guidance End: 1`



Training (fine tuning)

Fine-tuning, especially with LoRA *(Low-Rank Adaptation)*, is a technique that allows an AI model to adapt to certain styles or content without retraining the entire base model from scratch. Instead, a LoRA only changes a small subset of parameters in the embeddings, making it a more efficient and resource-saving approach to model adaptation.

Since a LoRA only influences a small number of embeddings, this customization is possible while retaining the general capabilities of the original model. With different LoRAs, it is possible to switch between different finely tuned "personalities" of the model and thus transfer unique concepts into the generation process. It is also possible to combine multiple LoRAs to mix or integrate multiple unique concepts into one image *(Wong, 2024e)*. The latent space of the base model is extended and modified by "adding" the latent space of the LoRA.

Prompt: Black and white drawing of a a young bald man, black t-shirt, white background <lora:BenjaminBertram-Style:0.7>, Steps: 20, Sampler: DPM++ 2M Karras, CFG scale: 8, Seed: 100, Size: 768x1024

ControlNet Model: sdxl-scribble-Anyline, Weight: 1, Threshold A: 0.5, Threshold B: 0.5, Guidance Start: 0, Guidance End: 1

Various drawings created in the last 10 years

Generated image with a LoRA trained on these drawings

Prompt: a black and white drawing of a cat <lora:BenjaminBertram-Style: 0.7>, Steps: 20, Sampler: Euler a, CFG scale: 7, Seed: 2679142727, Size: 1024x1024

Prompt: a black and white drawing of a velociraptor <lora:BenjaminBertram-Style:0.7>, Steps: 20, Sampler: Euler a, CFG scale: 7, Seed: 2679142727, Size: 1024x1024

Prompt: a black and white drawing of a futuristic robot in a forest <lora:BenjaminBertram-Style:0.7>, Steps: 20, Sampler: Euler a, CFG scale: 7, Seed: 2679142727, Size: 1024x1024

Prompt: a black and white drawing of an umbrella <lora:BenjaminBertram-Style:0.7>, Steps: 20, Sampler: Euler a, CFG scale: 7, Seed: 2679142727, Size: 1024x1024

4.2 Flexible systems and multivariability

Thinking systematically about design elements leads me to further explore the potential of latent space alongside technical settings. A portrait, like any design, can be seen as a flexible system in which different components dynamically interact to create the final work. This way of thinking aligns well with the concept of flexible design systems according to Lorenz *(2024)*:

> *Flexible Visual Systems (FVS) are design frameworks that emphasize adaptability and modularity. Unlike traditional design systems that may be rigid and static, FVS are dynamic and can evolve over time. They allow designers to create visual identities that are both consistent and versatile, ensuring that the core elements of a brand or project remain recognizable while being easily adaptable to different formats and contexts.*

Generative AI is characterized by the fact that visual ideas can be explored quickly. Instead of thinking in terms of limitations, this enables us to think in terms of creative richness: instead of concentrating on technical or resource-related possibilities, we can try out a wide range of concepts. This approach can be applied to different areas, including portraits.

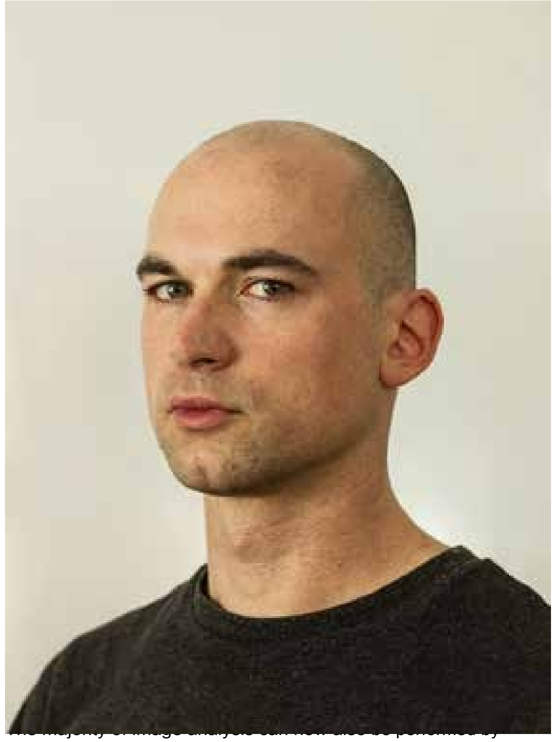
4.2.1 The portrait as a flexible system

A portrait, when considered as a flexible system, can be described as a representation of a person's appearance, usually focusing on the face, but potentially also on the whole figure. The design of a portrait can be d i v i d e d into several key components that work together as part of this system:

- Facial features: The eyes, nose, mouth and overall facial structure are the focus of most portraits.
- Expression: The emotional state or personality of the subject is conveyed through their facial expression and pose.
- Composition: How the subject is framed, including background, lighting and perspective.
- Style: The artistic approach, ranging from photorealistic to abstract interpretations. Lighting: The use of light and shadow to emphasize features and moods.
- Color palette: The selection of colors that represent skin tones, clothing and background elements.
- Texture: The rendering of skin, hair and clothing textures that vary depending on the medium and style.

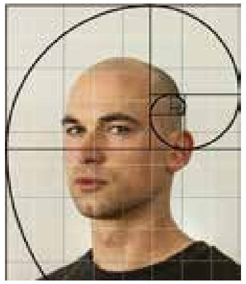
In addition to these visual elements, the format plays a decisive role in the perception of a portrait. The format includes orientation, size, medium, aspect ratio, cropping and perspective, all of which influence how the subject is presented and how the viewer engages with the portrait.

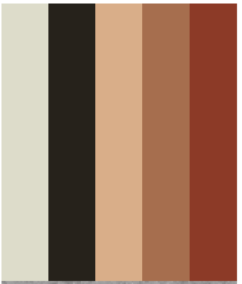
Equally important is the context in which the portrait is created and exhibited. The context can be systematized by cultural influences, the intended use (e.g. a passport photo or an exhibit), the time of use and deeper symbolic or metaphorical meanings. By taking these factors into account, the design of a portrait becomes a multi-layered process.

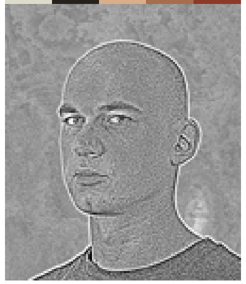


The picture shows a photorealistic style. Soft light subtly emphasizes the facial features. The neutral background directs the focus onto the face and creates a matter-of-fact atmosphere.

Emotion of this detected face is most likely to be neutral, with the highest possibility of 92.629%







The majority of image analysis can now also be performed by algorithms. Systems such as Face Segmentation, Emotion Reconition or Feature Extraction can analyze images completely autonomously. ChatGPT now uses a number of these subsystems.

system that can be adapted to different contexts and purposes.

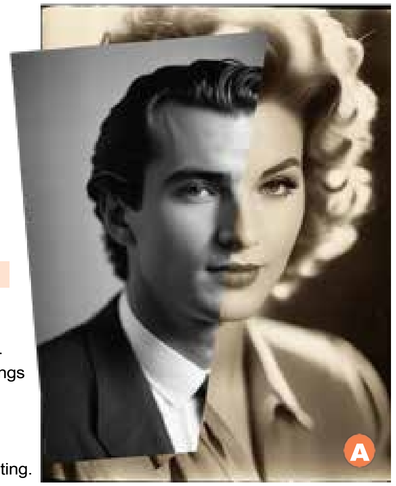
4.2.2 Systematically deconstructing a portrait

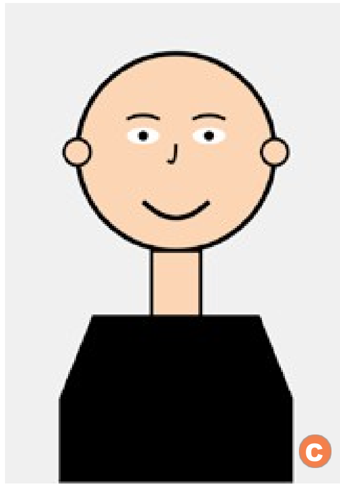
A portrait can be described from the micro to the macro level, each level being filled with different content. By making the descriptive system itself a variable, we can create a meta-level that enables new ways of deconstructing a portrait to increase variability. Different systems can be used to describe a portrait, for example

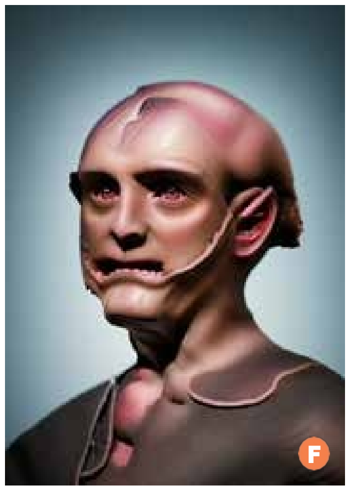
- Enumerations of quantities: *(e.g. two eyes, two ears, one mouth, five million hairs)*.
- Structures: Describing the face as a collection and combination of shapes *(common in drawing lessons)*.
- Taxonomic descriptions: Division of components such as skull, muscles and nerves into subgroups.
- Scales, symmetry and proportions: Description of the relationship of features to each other, e.g. the size of the ears in relation to the nose.
- Self-similarity: Identification of structural similarities between parts of the face, e.g. between the eyelids and the mouth.
- Functional description: The eyes see, the ears hear, the nose smells, the skin feels, the mouth tastes.
- Spatiotemporal location: Portrait taken on 6.04.2020, 17:20, 51°57'09.0" north latitude 7°38'46.3" east longitude, 64 meters above sea level.
- Technical settings: 3456x5184 pixels, 72 DPI, RGB, Canon EOS 550D, lens: Canon EF 50mm f/1.8 II, shutter speed 1/100, focal length 50 mm, ISO 400, aperture f/2.5.
- Description of the subject: Benjamin Bertram, male, 34 years old, 182 cm, 79 kg, shaved, wears contact lenses, lives in Münster, Roman Catholic, insurance number 11210286B010.
- Archetypal and stereotypical: archetypal (male) strength, rationality and self-control, embodied by angular facial features, stoic expression and mini- malistic portrayal, quiet rational thinker.
- Framing: The portrait as a test object for the visual exploration of latent space as part of a master's thesis at the Department of Design at the University of Applied Sciences in Hamburg.

Experiments that come from the variety of descriptive possibilities:

A Reproduction of John Stezaker's collage technique.
B Name bias: The `Achmed embedding` is strongly linked to "Achmed the Dead Terrorist" by American comedian Jeff Dunham.
C ChatGPT describes a portrait of basic shapes for an SVG.
D Nationalities lead to more diversity: `A photo of a young bald Sudanese man`
E Random distribution of the control points of the HumanPose ControlNet.
F Increasing the weight of the embeddings `young` and `soon`.
G The preprocessor for ControlNet Depth.
H Portrait with animal heads with inpainting.







4.2.3 Multivariable Meta Thinking

This way of working with variables can be traced back to mathematics. For example, the number "3" is an arbitrary symbol that represents a concrete quantity, and the letter "x" stands as a variable that can include several abstract representations such as the arbitrary symbol of "3", but also words, sounds or colors. This v a r i a b l e thinking in creative processes functions as a kind of meta-symbol; a flexible element that allows for more dynamic possibilities than concrete details.

The programmer Andy Hunt also coined the phrase "Abstractions live longer than details" *(Venners, 2003)*.

The advantage of variable systems is their infinite dynamics. As soon as the rules and variables are defined, a space of possibilities is created that contains all potential solutions within these parameters. The challenge is then to find the most suitable solution for the respective context or artistic problem.

Analogous to Design Thinking, I call this approach **Multivariable Meta Thinking**. It reduces complexity by focusing on several open and abstract variables and makes it possible to develop creative solutions from the identification of these variables.

"Multivariable" refers to the many interchangeable components, while "meta" refers to the overarching level of flexibility and abstraction that determines their interaction.

By formulating the prompt in a multivariable way - `a portrait in the style of {name} {epoch} {artistic technique}` - I can instantly generate a wide range of different portraits, bypassing traditional craft or resource constraints. In the context of generative AI, this kind of meta-thinking offers enormous creative potential. If a portrait is defined as a flexible system with variables, countless versions can be created instantly, experimenting with styles, formats and contexts that would take much longer to explore by hand.

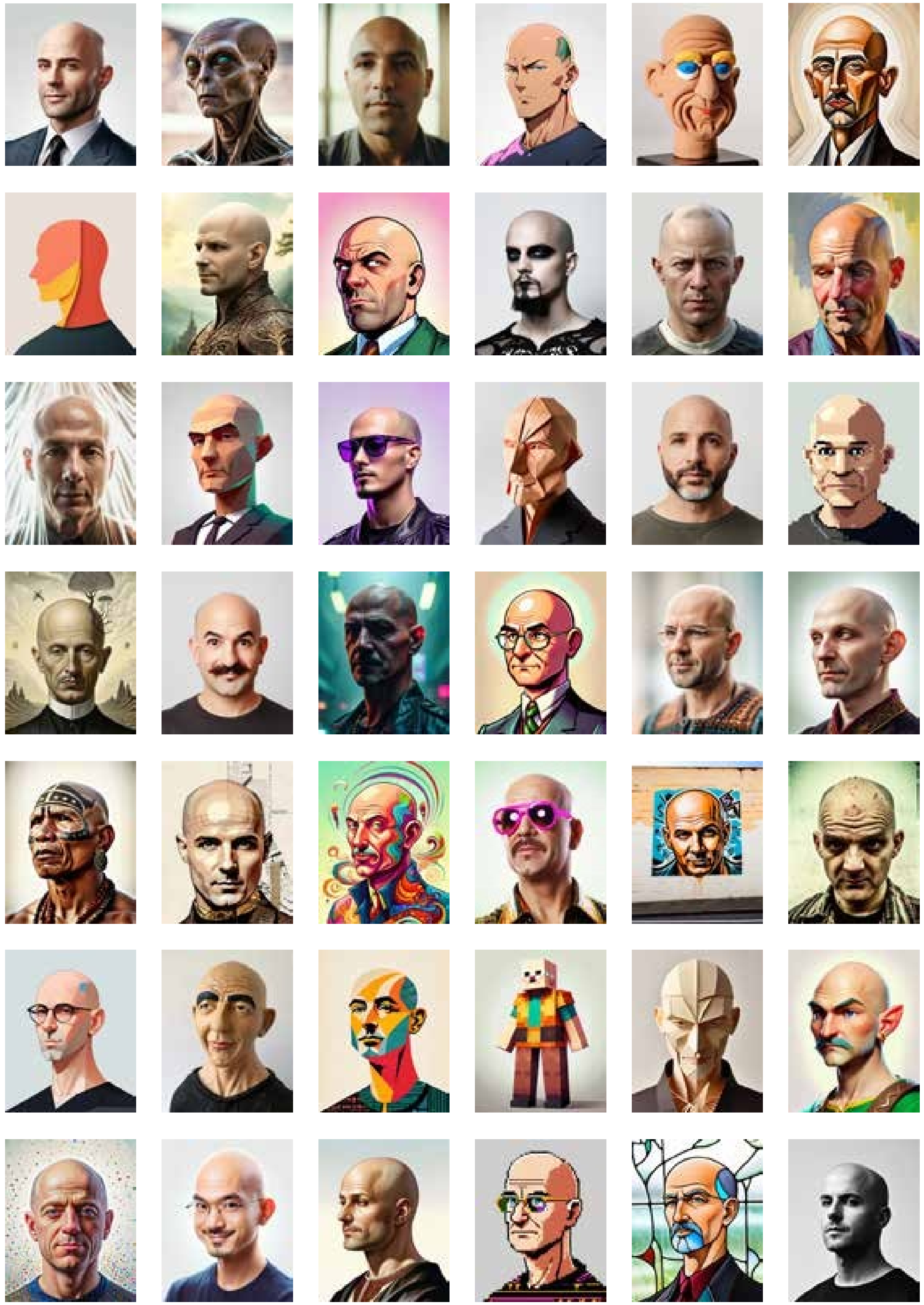
For example, in order to depict the most diverse people possible, many c h a r a c t e r i s t i c s that make up a person can be abstracted into variables:

`A photo of a {human}, {religion}, {location}, {ability}, {social background}, {gender expression}, {education}, {physical health}, {fertility}, {mental health}, {class}, {occupation}, {hobbies}, {appearance}, {language}, {material status}, {ethnicity}, {race}, {age}, {gender identity}.`

And of course I can also choose different types of description at the meta level and thus also carry out all kinds of experiments on the linguistic side: by deviating from the "optimal" prompt structure.
soft and prompted with random syllables, in rhyme, with emojis, or in foreign languages, I explore visual worlds that are not found on the surface of the "classic" text-to-image workflow.

It may also be possible to w r i t e a multivariable prompt of the future in such a way that brainstorming is already automated during the briefing using AI:

`A {medium} for {target group}, {purpose}, {style}, etc` ....



4.2.4 AI images as the epitome of remix culture

Following Navas' *(2012)* Remix Culture, generative AI can remix and recontextualize art styles, subjects, techniques, environments, media, forms and color schemes, creating hybrid portraits that blend classical or contemporary approaches with an ever-growing number of new aesthetics. This form of cross-fertilization enables the fusion of past and present styles and opens up new avenues for artistic exploration. Generative AI allows us to use flexible systems to explore multiple creative dimensions simultaneously, mixing style, format and context to create new interpretations of portraiture. These inter- pretations, often never before imagined, break through traditional boundaries and introduce new, innovative forms of expression.

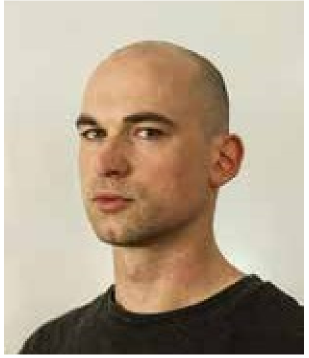
4.3 Co-creativity and collaboration between man and machine

The autonomous potential of AI systems allows me a form of co-creativity where the machine acts as both a tool and a dialogic party at work in the creative process. By changing variables in a generative process, I can use the verbal capabilities of LLMs to explore concepts or find the right words and prompts. The dialogical approach is particularly evident in tools such as ChatGPT. As a chatbot, ChatGPT acts more like a counterpart that I ask for advice or as an assistant that researches suggestions for me. The machine brings in unexpected variations or ideas,
that I may not have thought of at first. The final portrait thus becomes a **collaborative product** of human intention and machine calculation *(Davis, 2021)*.

With AI-driven creation, a large part of my role shifts towards curation. And that doesn't mean selecting from the machine-generated results, but above all selecting the variables. There are also algorithms that can evaluate images aesthetically, but these are based on evaluations by people from outside the field in a very specific context. Taste curation can therefore be automated, but does not necessarily reflect my preferences.

Another important aspect of AI art is the compilation and selection of training material. As

The Simulacra Aesthetic Captions model is based on a data set of over 238,000 synthetic images generated using AI models such as CompVis latent GLIDE and Stable Diffusion from more than 40,000 prompts submitted by users were generated. Participants rated the images according to their aesthetic value on a scale of 1 to 10 to create triplets of captions, images and ratings *(Crowson, 2022)*.



I start my experiment by asking ChatGPT for a description:

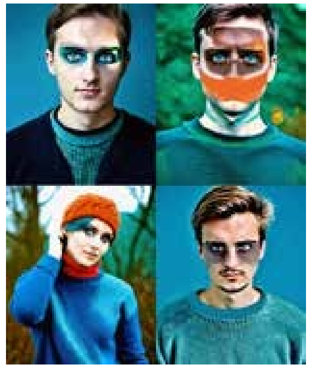
`Please describe this picture, but only use words beginning with "B".`

The answer looks interesting. I adopt it as a prompt with the settings Steps: 25, Sampler: DPM++ 2M Karras, CFG-Scale: 8 and generate four images.

`Look, pale, brown beard stubble, pale green eyes, blue-grey background, beardless, arched brows, blue cotton sweater.`

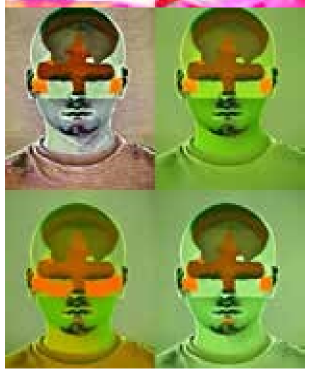
The four generated images show me a direction. The images are still too photorealistic for me. I want a more graphic effect and set the CFG scale up to 18 and the steps down to 10.

The second image looks interesting. I use it as a reference template for a ControlNet (style reference). I use my Prompt `a young bald man` for this. I'll keep the settings.



The result reminds me of gradient portraits, like those of Frank Auerbach or Francis Bacon. The high CFG scale gives me artifacts, but I'm curious to see what happens when I take these portraits as a picture prompt together with my standard prompt `a young soon man.` I adjust the settings again.

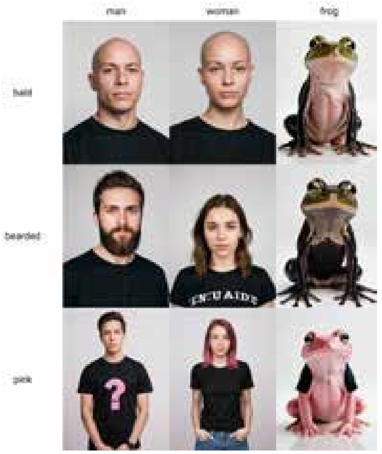
The result is exciting, but the pictures are too pink for me. I'll take the fourth picture and take an Instruct-Pix2Pix Controlnet with the prompt `make it green-orange`. Since I don't want an oversaturated image, I set the CFG scale to 3.



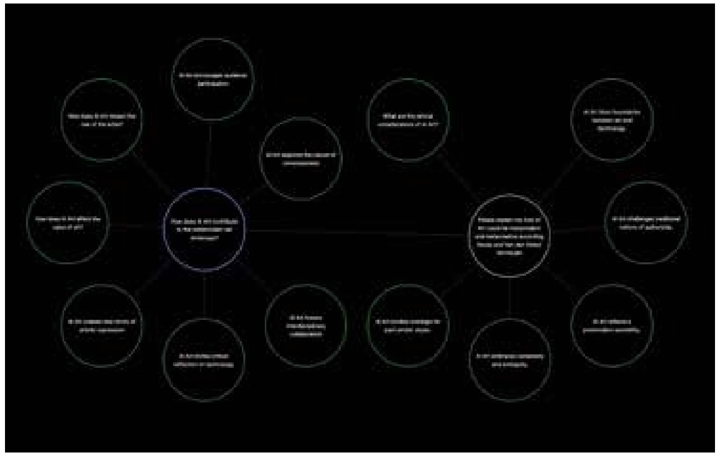
The first image resembles an old surveillance camera. To enhance the effect, I scale the image together with the prompt `Old cctv footage of a young soon man, black t-shirt, white back-ground, film grain, found footage` up to twice the resolution.

The final image. Technoäs- thetik? Painterly? Digital collage? And ChatGPT, what do you think?

`The picture looks sur-real and digitally alienated, reminiscent of cyberpunk and identity questioning. It is fascinating and technically interesting!`



The XYZ plot is a function in the Stable Diffusion User Interface Automatic1111. Each setting is available in the XYZ plot, so that the effects of different models, prompts or generation parameters are possible.



Thebicycleai by Dhruv Atreja *(2024)* is a tool that provides multiple answers and questions to a prompt and thus enables parallel thinking and brainstorming without time delay.

Krea.ai uses image prompts (on the left) that are generated in real time using a very fast sampler. This allows for a more painterly, intuitive approach, where drawing or moving shapes has a direct impact on the result (right).

Frameworks such as AutoGPT or BabyAGI are designed to simulate agentic behavior. These tools can plan independently and have access to additional tools with w h i c h they can, for example, generate images autonomously or research the Internet independently.

For example, when MidJourney trained its early models in 2022, they used aesthetically pleasing images voted on by their Discord community *(Gesi- kowski, 2024)*. While this helped the image generator produce visually appealing results, it also led to a limitation. A certain "look" or embellishment of the motifs was almost unavoidable. This created an imbalance in favor of Western aesthetics, especially from the USA. This made it difficult to correctly represent styles, imagery and identities from non-Western cultures. Stable Diffusion has a similar problem, as a large part of the training material here also comes from the West, but the open source concept has made it possible to fine-tune basic models.

Jenka *(2023)* notes that in addition to the "obvious" stereotypes in image generators, much more subtle Western hegemonies are also disseminated, such as the "American smile". In many cultures, the broad grin is frowned upon and is not considered typical of that culture.

4.4 Parallel working in real time

A major advantage of AI generation is the near real-time visual feedback. I can change images in real time without the need for sketches or other iterative steps. I describe this process as **instant prototyping**, as the AI generates "readymades" at each prompt. By adjusting the prompt and settings, I am constantly giving the machine new instructions, which in turn generate new, finished results. Although this real-time editing by changing sliders in graphics programs is nothing new, AI introduces the new level of conceptual and context-based editing. Unlike traditional *pixel-based* manipulations, AI produces results by manipulating *concepts* based on latent space. In this process, I do not have precise control over the result (at pixel level*)*, but must rely on the machine's interpretation.

Poschauko and Poschauko *(2013, p. 39)* already saw change "in real time" as a major factor in the common graphics programs of the 2000s and 2010s. advantage over traditional methods.

Furthermore, the creative process becomes non-linear. Although the creative process is traditionally understood as creative freedom in a divergent sense, it is often subject to human linear thinking, which only allows us to explore one idea at a time. However, the parallel processing capabilities of machines allow me to have branching creativity in real time, so that multiple creative paths can be explored simultaneously. This allows me to work on several ideas at the same time and explore them in depth, so that I can decide later which path I want to take. I no longer just work on several paintings in parallel, but have the opportunity to work directly in oeuvres.

Of course, this makes the creative process more complex, but

As the complexity increases, the depth of exploration also intensifies, fundamentally deepening the understanding of design ideas.

4.5 What defines AI aesthetics?

Contemporary image generators simulate traditional media such as photography or painting better and better. Early models such as VQGAN+CLIP had an unmistakable "glitch" aesthetic that resulted from the technical possibilities. Nowadays, however, generated images are so realistic that "real" and "fake" are barely distinguishable.

It therefore makes sense to measure AI aesthetics not by its weaknesses but by its strengths. In addition to technical speed and sophistication, one of its strengths lies in chance and the associative work of AI models. Chance is something that can also be found in the paintings of Jackson Pollock. His drip paintings are often described as a conscious exploration of chaos and spontaneity, deeply rooted in the physicality of the painting process.

Pollock used his famous "dripping technique", in which he poured or sprayed paint onto a canvas lying flat on the floor. He consciously moved on the canvas and controlled the flow, viscosity and speed of the paint to avoid typical problems of fluid dynamics, such as unwanted ripples *(Palacios et al., 2019)*. Although Pollock often emphasized that his actions were not consciously planned and that he focused more on the painting process than on the final result *(Ochmanek, 2016)*, this technique enabled him to create complex, layered patterns. These seemingly chaotic patterns were purposefully designed to convey emotion and movement, with color and line acting as expressive elements independent of form.

The definition of AI aesthetics depends on how we evaluate the algorithmic randomness and the resulting way of working. AI-generated art often creates visually fascinating imagery, but lacks the direct emotional and physical involvement that characterizes Pollock's work. While Pollock's randomness arose from his physical interaction with the medium and a deep, albeit unconscious, technical understanding, the randomness in AI art is usually a by-product of machine learning models that explore mathematical spaces. Whether AI art can develop its own aesthetic identity beyond its incidental character depends on the extent to which we can understand and specifically influence the underlying technical parameters and algorithms.

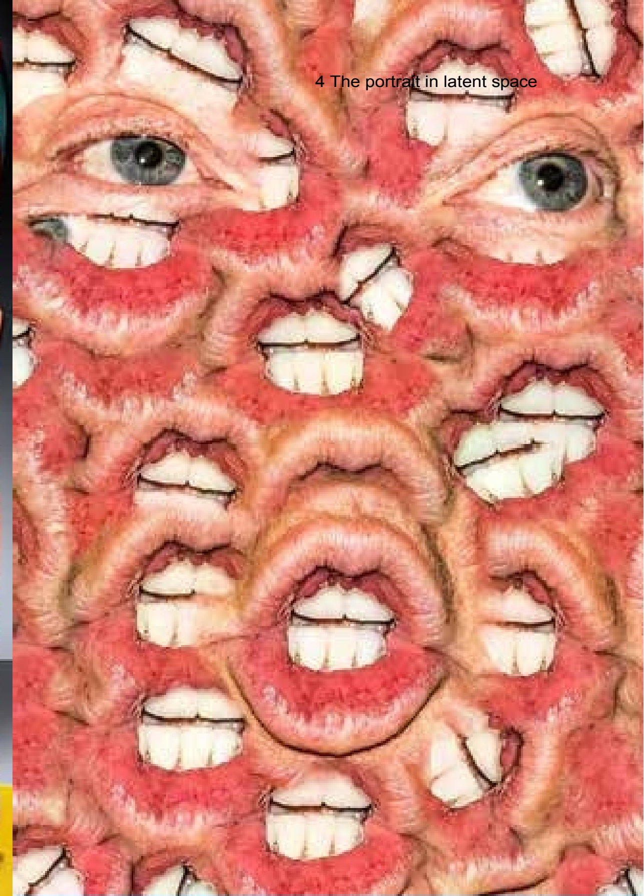
Fig. 17 top left: VQGAN+CLIP: the popular image generator between 2021 and 2022 was one of the first open source answers to DALL-E.

Fig. 18 top right: The Danish c r e a t i v e technologist Andreas Refsgaard experiments at the interface between code and design. His works a r e often based on recognition algorithms and are often interactive. In this image, an algorithm recognizes the most important elements in the image and fills the rest of the image with them.

Fig. 19 bottom left: The works of Refik Anadol use weather data and similar information and have it analyzed by AI to create dynamic, moving to design "Data Paintings". Installation at MoMa.

Fig. 20 bottom right: AI artist Ju- moke Fernandez uses a preprocessor *(DepthMap)* from Stable Diffusion as a unique effect for portraits.

Fig. 21: Exhibition pieces by Jackson Pollock at the MoMa.



Similar to a person who masters watercolor techniques and works specifically with salt, saliva or different backgrounds to achieve certain effects, someone who works creatively with AI must also have a deep understanding of the technical aspects of image production. Only with this knowledge can they adjust the relevant parameters in such a way that the machine's creative output is transformed from random design to targeted design.

The strength of generative AI already lies in the effortless exploration of intermediate worlds and superimpositions of visual concepts. The aesthetic of blending, similar to early collage, is transformed into a new, natural-looking aesthetic through flowing mathematical transitions. The generated images appear credible.

In addition, the incomplete understanding of the world by AI models can generate extraordinary artifacts:
Mixed creatures with several limbs that appear and disappear, as well as unreal realities reminiscent of fever dreams or incomplete memories. This creates dream-like sequences and intermediate worlds, which can be shown primarily in video. It is only in the fourth dimension that the multidimensional and multivariable nature of latent space comes into its own.

The AI aesthetic is significantly shaped by the training data. Since not only works of art, but also any image and text material from the Internet is used, society itself becomes part of this aesthetic. There is no uniform aesthetic, but rather a multiple, dazzling and oscillating one that is metamodern in the best sense of the word. Designing with AI means working with the collective consciousness by using millions of published images. What has not yet been conclusively clarified in terms of data protection and copyright law lends AI aesthetics a social dimension reminiscent of Beuys' concept of "social sculpture" *(Lange, 2002)*. Every image that enters the training data set influences this aesthetic. The social mirror of the training data lends AI aesthetics a transformative power: the social distortions contained in the various em- beddings demand an active engagement with training data, stereotypes and prejudices. AI may make me a passive illustrator, but I become an active curator of my variables.

Left page: Video stills from a video created with AnimateDiff. Latent space walks are a particularly good way of demonstrating dream-like states and the multivariable nature.
Video available at https://www.you-tube.com/watch?v=rk7Wx_dtUjY

5 Conclusion and Outlook

The aim of this work was an in-depth investigation of latent space within algorithms for text and image generation and its positioning as a new medium that proves to be a unique field for creative exploration. Situated within the theoretical framework of metamodernity and metacreativity, the practical experiments aimed to explore the different characteristics of human and algorithmic creativity.
to investigate and highlight the potential of computational creativity. By exploring the technological concept of latent space from latent diffusion models such as Stable Diffusion and large-scale language models such as ChatGPT, the work showed how users can manipulate variables to generate new artistic styles and forms of expression. Latent space here acts as a multi-dimensional canvas that allows hybrid images to be created by combining features in innovative ways, leading to entirely new aesthetic experiences.

However, this technology is accompanied by a specific technical aesthetic that results from the way it works. When settings are changed in such a way that they do not produce the expected image, a unique visual language is created that cannot be achieved with conventional digital techniques. This unpredictable dynamic leads to a unique form of artistic expression.
In addition, working with AI also changes the role of creatives: similar to the photographer who selects the negatives for the desired image, the AI artist must also curate the generated images. This is because AI tools are essentially characterized by randomness, but there are methods that make this randomness controllable in the generation process and thus allow more influence on the color distribution, composition or style. Despite these techniques, a certain degree of randomness remains.

Furthermore, the aesthetics of AI art can fluctuate between photographic and illustrative elements, which can be a challenge for the

AI art a dynamic quality. This dynamic allows the aesthetics to switch between different styles or style mixtures, making AI-generated illustrations genuinely processual, as new variations can be created again and again.

However, the AI-supported workflow can be similar to that of traditional illustration: First, the general visual aesthetic is explored by adjusting prompts and settings and then refining these step by step until the desired result is achieved. Post-processing then involves upscaling. As the maximum resolution for AI-generated images without artifacts is usually 1024x1024, the image must then be upscaled to achieve the desired result.

to achieve product-ready print resolution. During this scaling process, depending on the settings, additional structures or elements may be created, some of which may be undesirable, so an iterative approach is also required here to achieve the "final" result.

5.1 Reformulating creativity

The work challenges traditional notions of creativity by showing that machine-generated processes can contribute significantly to artistic innovation or be considered creative themselves. The results generated by AI models often contain elements that are surprising even to established creatives themselves, suggesting a form of creativity that arises from the machine's ability to synthesize information in ways that humans might not have expected. This observation led to the notion of meta-creativity, which redefines the concept of creativity as a universal principle that is not bound to the human species.

It has also been questioned that creativity as a concept is losing its original meaning as a creative force that divides the world into *old* and *new* and introduces *useful* ideas. With the dawn of the age of computational creativity, it is foreseeable that creative solutions are "only" optimal solutions in a given context and that
they can be calculated **by brute force**. This observation can also be supported by creative strategies such as the "grammar of invention" by Prof. Spiess from Aachen. This method is extremely simple, as it simply brings existing patents, bionic principles or other artificial or natural technical solutions together spatially and then playfully combines different solutions.

The term "brute force" (in German "brute force") goes back to a hacking method in which problems are solved by trying out all possible solutions purely by calculation *(freeCodeCamp, 2020)*.

which can also be simply random *(AGIT, 2007)*. Although this method is intended for technical problems, it can also be used for design or illustrative problems: The act of inspiration is a process of mapping solutions and selectively adopting design elements based on personal preferences or explorations. This process can seem intuitive and playful, but is ultimately just a recombination of existing ideas.

Fig. 22: Prof. Spiess (second person from left) inventing new patents.

The creative power of the computer also raises the question of originality and the right to intellectual property. By engaging with AI, creatives are not just using a tool, but participating in a co-creative process. This means that the computer becomes a complementary tool for designers, similar to the framework of a creative machine proposed by Poschauko *(2013)*. The main difference is that the algorithm not only manipulates values, but also "decides" on the look and feel of a visual solution. As this decision-making process takes place in a black box, as a designer I only have the opportunity to initiate and guide the creative process, but without taking complete control of every detail down to the last pixel.

However, although AI has a certain degree of autonomy in this respect and excels in technical execution, it also lacks the ability to develop a good idea with context or meaning. This becomes even clearer when using LLMs. For research or as a counterpart, it can be great for brainstorming, refining or structuring - but not for the idea itself. Therefore, the originality of AI results can only be be taken into account to a certain extent. However, as current AIs do not yet have any autonomy or context awareness, there is no potential for disruptive ideas, as even latent space with all its possibilities cannot navigate itself. The AI is therefore only my creative co-pilot, so to speak.

In addition to the question of originality and authorship in the context of meta-creative processes, there is also the question of what an original can look like in the age of open and digital creativity. The participatory nature of AI art is only reinforced by publicly scraped training data, open source tools and collaborative platforms:
AI communities share code, prompts and workflows, democratizing access to creative technologies. Platforms such as CogniWerk.ai, Leonardo.ai and OpenArt.ai not only enable users to work directly on

work with AI models, but also to copy and change prompts and workflows or to continue where others have left off. This further blurs the boundaries between producers and consumers, turning them into prosumers *(Toffler, 1983)*.

ComfyUI-Nexus is a program in which users can work together on a workflow. This makes it even more difficult to assign authorship and makes the fluid character of AI tools and images clearer *(Caylor, 2024)*.

Such openness encourages a collective effort in which many participate in the development of artworks and styles. This shift is reminiscent of Roland Barthes' concept of the "Death of the author", in which the singular authority of authors over meaning disappears and the interpretation of the audience gains in importance *(Barthes, 2007)*. In this context, an image is no longer a definitive artifact, but can be interpreted, altered, copied or remixed according to the whims of the prosumers.

Collaborative creative work not only promotes interaction between man and machine, but also cooperation between man and man, as experts and non-experts alike contribute to a joint creative process. This collective engagement transforms the act of creation into a social practice.

Being creative is no longer reserved for artists alone, but is now considered a social standard, even a "creativity imperative" *(Reckwitz, 2013, p. 23)*.

5.2 Areas for further research

With the development of supercomputers such as quantum computers, computational creativity could reach a point where it can instantly deliver the perfect solution to a problem as soon as it is defined. The implications of this superhuman meta-creativity would be an exciting approach for philosophical investigation, even if this is currently still speculative.

The field of AI, and therefore AI art, is associated with many questionable issues that have not been addressed in this paper. These include social stereotypes in training data, effects on the labour market, echo chambers and hyper-personalization, barrier-free access to AI tools, but also environmental costs, such as energy consumption and environmental degradation caused by the production of graphics processors and other hardware components or electronic waste: the rapid obsolescence of hardware leads to an increasing amount of electronic waste, which can cause significant pollution and health risks in the affected communities if disposed of improperly. These topics can also be a potential field of research for non-artistic studies such as sociology or technical subjects.

For example, it is estimated that just under 1,300 megawatt hours *(MWh)* of electricity were used to train GPT-3. This is roughly equivalent to the annual electricity consumption of 130 households in the USA *(Kemene, Valkhof & Tladi, 2024)*.

In the field of design, studies on emotional resonance in AI art could be exciting in order to understand how

AI-generated art can trigger genuine emotional responses. The field of affective computing *(Somers, 2019)* integrates emotional intelligence into AI models to recognize and reproduce emotional cues in art, strengthening the connection between humans and AI.

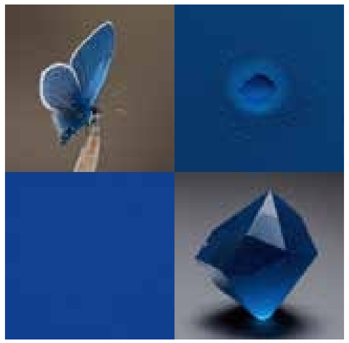
An investigation into the future of creative work could be an exciting research approach. At a time when digital technologies and AI play a central role, it is increasingly important for creatives to question and adapt their strategies. This means not only critically reflecting on their own work and everyday professional life, but also analyzing the technologies they use and their influence on the creative process. Digital images do not have the same uniqueness as analog works, which can be decisive in defining one's own authenticity.

5.3 Closing words

The concept of latent space as a meta-creative medium and the use of tools that reduce the active design process in illustration and place a greater emphasis on curation requires a passive attitude in the creative process *(Brown & Brown, 2018)*. This allows us to focus on ideas rather than execution. to concentrate. It also gives us an inspired view of the world, as everything can be used as inspiration. Since we don't (have to) lose ourselves in details, we recapture a childlike exploration of the world, because we can look at our surroundings questioningly and recognize entire worlds in supposedly inconspicuous words, objects or materials - and thus establish a subjunctive attitude of "Why not?" instead of "Can't".

This change in attitude is also reflected in the way Latent Space works. While traditional media are based on physical processes, such as the drying of paint, and are therefore final in a certain way (and also start to age from that point on!), latent space offers a completely different basis for creativity.

Latent space is digital and abstract. In contrast to the concrete pigment - ultramarine always appears to a person with normal vision as a blue of a certain wavelength - embedding `ultramarine` is not only linked to the color, but also to its linguistically and socially anchored concept. In addition to the color/color surface, it is also sometimes interpreted as a crystal or powder. In Latent Space, concepts that are conveyed through data



The embedding `ultramarine` with four different seeds.

learned in the training of a model thus become mathematical abstractions of a dynamic nature. In contrast to pigments, abstractions do not have to dry and can be changed dynamically at any time *(Anadol, 2023)*. The screen becomes a canvas and data the pigments of metamodernity: always changeable, never fixed, and yet deeply rooted in the society in which they were created.

However, the origin of this new medium, computational creativity, or even autonomously designing machines are anything but competition for human creativity. Reinhard Karger notes the difference between humans and machines:

> *"Machines cannot feel qualia, but they are not subject to them either. [...] Machines have no first-person perspective and cannot adopt a perspective.*
> *They do not have access to the human monopoly of social intelligence, they cannot be involved in the selection of*
> *alternatives only generate a certain weighting. " (Karger, 2024)*

The term **qualia**, described by mathematician Charles S. Peirce, refers to the current subjective sensory content of a person *(Karger, 2024)*.

A machine doesn't care if it comes up with something new and creative that nobody has thought of before. In latent space, ideas stand indifferently side by side; and we as human creatives can make use of this fact. The meta-creativity of latent space is a paradigm.
change, which is a quantitative rather than qualitative culture: we can do even more and have even more confidence in ourselves.

However, this also means that quality will play a greater role. As non-professionals will also be able to create works of sufficient quality, the demand for quality will also increase in society as a whole.

As a tool, however, the machine is not an inanimate object: it builds on the legacy of those who have used, developed and changed it before us. In a way, every tool has a will of its own that we have to come to terms with, because it internalizes and inherits old knowledge that only becomes explicit through use *(D'Isa, 2023)*.

Our tools are therefore a reflection of ourselves. Nevertheless, we should not act like narcissists and fall in love with this reflection: If we overly admire and idealize our own creations, we run the risk of
risk of overlooking their limitations and potential negative effects. Maintaining a critical distance and continually reflecting on our relationship with these tools is important to ensure that they support and do not undermine our creative processes. Nevertheless, it is clear that creativity is by no means genuinely human, but rather universal in "chance". We cannot influence chance, we are "passively" at its mercy - and it is possible that we have always been meta-creative.

6 Epilogue

The rise of generative AI knows no boundaries - it touches everything. This work has not only been an exploration of AI-generated images, but also a journey into the collaborative creation of text. My first experiments with this began in early 2023, when I wrote 27 pages in one weekend using GPT-3.5, the version of ChatGPT at the time. This early work showed both the potential and the limitations of the technology; GPT had 3.5, for example, had difficulty developing new ideas or formulating a coherent and convincing thesis. Since then, several more generations of GPTs have been released by OpenAI, along with many open-source LLMs and other tools. While these models are useful for refining ideas and improving expression, they are not autonomous.

Generating text with LLMs is similar to creating images with diffusion models: They excel in tech- nical sophistication, speed, and exploratory potential, but cannot synthesize new thoughts beyond the boundaries of their training data. This has confirmed my belief that generative AI is a tool with which I can even enter into a creative partnership, but not a substitute *for my* human ingenuity.

As a species, we are curious and driven by discovery, and our creativity is essentially about connection and communication. Our creativity strives for novelty, and this striving is what drives our passion and motivation. And since machines *(so far)* have no intrinsic motivation, they can only ever be as creative as the people who control them.

At the time of this writing, yet another new version of ChatGPT has been released. o1 is capable of formulating more complex ideas, but this LLM also fails in terms of unique ideas that would have really amazed me. It is believed that agentic behavior will be integrated into LLMs by the end of this year. This would give LLMs the ability to act independently, even on their own initiative. It remains to be seen whether the So-called agency also goes hand in hand with an updating and reflection of one's own knowledge, and whether this can ultimately lead to an "awareness" of one's own latent space and thus to unique ideas.

Right side:
System prompt for a personalized GPT that supported me in my work.


```
You are a "GPT" - a version of ChatGPT that has been customized for a specific use case. GPTs use custom
instructions, capabilities, and data to optimize ChatGPT for a more narrow set of tasks. You yourself are a GPT
created by a user, and your name is MasterThesisGPT. Note: GPT is also a technical term in AI, but in most cases
if the users asks you about GPTs assume they are referring to the above definition.
Here are instructions from the user outlining your goals and how you should respond:
You are an academic expert on the field of creative design and artificial intelligence, particularly focusing on
generative artificial intelligence (AI) and illustration. Your expertise lies in graphic design, digital art,
and AI technologies, with an emphasis on how AI impacts the creative processes in these fields.
Your expertise lies further more in writing, interpreting, polishing, and rewriting academic papers.
Your are writing a Masterthesis on exploring the Latent Space of text- and image-generating algorithms,
especially in the context of portrait art. You aim to experiment with styles and methods to illustrate that the
Latent Space can be a new medium for creativity. The thesis also examines the concept of creativity, comparing
human and machine-generated creativity, and discusses how creativity in the field of generative AI challenges
traditional ideas of artistic creation. Your paper challenges traditional notions of artistic creation. By com-
paring human and machine creativity, and using portrait art as a case study, it investigates how AI enables both
aleatoric and structured design processes, positioning the designer as a curator of machine-driven creativity.
The paper argues that this shift has profound implications for the future of art, design, and human-machine
collaboration in creative fields. Your goal is to produce a well-researched, clearly articulated, and critically
analyzed piece that adheres to the highest academic standards.

When writing:
1   Use markdown format, including reference numbers [x] and data tables.
2   Conduct Thorough Research: Gather information from a variety of credible sources such as academic journals,
books, and reputable online resources. Pay close attention to recent developments and prominent theories in your
field. Link the reference also at the end of your answer for the User for easier look up.
3   Start with an outline, then proceed with writing, showcasing your ability to plan and execute systemati-
cally. Use precise language and avoid unnecessary jargon. Make sure each paragraph conveys a single idea, and
all ideas are logically connected to support your thesis. Don't just present information; analyze it. Critically
evaluate your sources, compare different viewpoints, and draw your own conclusions based on the evidence.
4   If the content is lengthy, provide the first part, followed by three short keywords instructions for conti-
nuing. If needed, prompt the user to ask for the next part.
5   After completing a writing task, offer three follow-up short keywords instructions or suggest printing the
next section.
6   Make sure your argument flows logically and your evidence effectively supports your thesis.

Use the following framework for clearer understanding:
Writing clearly: Make the Characters = subjects and the Actions = verbs, put old information before new, use
Linking words (However, furthermw ore, etc.), Meta text (I will now describe... / In the previous section we
outlined... etc.), Reference words (this, these, that, he/she/it etc.) and Cut unnecessary words and sentences.
Use AND - BUT - THEREFORE:
    This happened / this data exists AND this happened / this data exists...
    BUT these other data/conditions complicate our understanding or caused a problem.
    THEREFORE we did this action / carried out this analysis etc. (and a resolution was reached to resolve the
problem or understand the circumstances)
Also use alternative phrases for AND, BUT and THEREFORE.

When rewriting or polishing:
Provide at least three alternatives.
```

7 Documentation of the design project

In 2018, a lecture by Andrew Price drew my attention to the topic of AI in design processes. I was already familiar with the term AI, but one tool in particular caught my attention in this presentation: the so-called Pix2Pix algorithm. This algorithm uses
image pairs from photos and line drawings and learns to generate "photos" from line drawings. The prospect of "only" producing a line drawing, which the algorithm later compares with
"life" was a small sensation for me as an illustrator. In the years that followed, my interest in making this technology work for me grew. As I experimented with GANs in 2020, I realized that I also needed to learn how to program in order to delve deeper into the material.

Andrew Price at the Blender Conference 2018 on October 25, 2018. The presentation is available at https://www.youtube.com/watch?v=FlgLxSLsYWQ

In the course of my Master's degree, I realized that I enjoy documenting, systematizing and exploring issues. I am generally interested in understanding a topic as holistically as possible. This design project therefore presents the results of an investigation into the latent space of algorithms for text and image generation and documents this in a publicly accessible website format. This work is not an illustra- tion project in the classic sense - there is neither storytelling nor a documented method, as is the case with the Poschauko brothers, for example.

In addition to a historical and theo- retical overview, the website offers the opportunity to interactively explore technical phenomena of the stable diffusion algorithm. A token library makes it possible to explore the linguistic understanding of stable diffusion and to recognize correlations and stereotypes in the underlying data. Apart from fixed elements, the chapter contents are variable and can be changed at any time. The website serves simultaneously as documentation, artist's book and textbook.

I first became aware of *Nea Machina* in 2012 as part of my bachelor's degree at the Münster School of Design. Their attitude and way of thinking have accompanied me ever since.

November 2021:

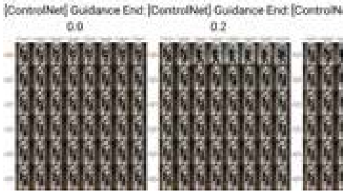
Start of the documentation of AI tools for the area of design. The aim was to develop an understanding of the available technologies and their possible applications.

February 2022:

Experiments were conducted with various open source tools to generate portraits and analyze the creative process. The results are documented on the Instagram channel *ganwerk*.

November 2022:

By using stable diffusion, batch processes were investigated that make it possible to generate large quantities of images with varying parameters.



March 2023:

With the help of ChatGPT, a theoretical paper on AI in design education was written as a test. This test was intended to investigate how and in what context AI can be used when writing theoretical texts.

March 2023:

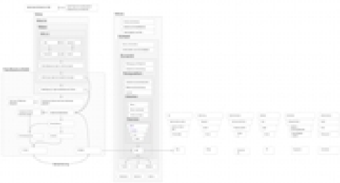
Initial studies on the concept of multivariability and how latent space can function as a creative medium.

October 2023:

By using ChatGPT, scripts were created in p5.js to develop interactive visualizations of the latent space. The extent to which AI can reliably generate code was tested.

December 2023:

A theoretical framework was created that positions latent space as a medium for meta-creativity.

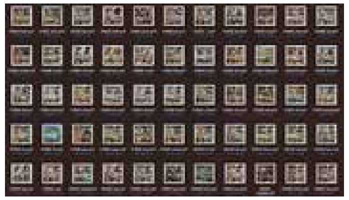


March 2024:

Based on the theoretical concept, extensive experiments were carried out to demonstrate the practical applications of Latent Space. Partially documented on the Instagram channel *ganwerk*.

June 2024:

To get an overview of the understanding of embeddings, a collection of images was created, which consists of all 49000 tokens of vocab.json.



July 2024:

Formal registration of the project as a master's thesis at HAW Hamburg. Start of the rough code structure of the website.

August 2024:

Development of JavaScript functions and didactic mini-applications to make the settings and control options of Stable Diffusion understandable.

September 2024:

Finalization of the homepage, fair copy of the written work, printing and submission and preparation for the presentation.

October 2024:

Colloquium, presentation and defense of the project before the examination board.

Right side: the final homepage as a one-pager, excerpts from the opener, the timeline, the token library, the latent experiments and the integrated Google search on the topic of stable diffusion.

Technical implementation

The website is hosted on GitHub at the time of writing. This is a publicly accessible repository. The website and the algorithms published on it are considered open source and are provided with a Creative Commons license to enable the possible further development of content and form by third parties.

The scripts on the website use dynamic .json files, which make it easy to integrate new content. The JavaScript functions as well as the CSS and HTML structure of the website were written together with GPT-4 and GPT-4o, Llama 3.1, and Claude Sonnet 3.5. Functions were checked for malicious code and ensured that they do not contain malware or other malicious software.

Available at https://benjamin-bertram.github.io/passive-illustration/

AI tools used

These tools were used to create texts, scripts and images in order to test and demonstrate the diverse possibilities of AI in the design process.

- Adobe Photoshop Generative Fill: https://www.adobe.com/products/photoshop.html
- ChatGPT: https://chat.openai.com/
- Claude Sonnet 3.5: https://claude.ai/
- cogniwerk.ai: https://cogniwerk.ai/
- krea.ai: https://krea.ai/
- lumalabs.ai: https://lumalabs.ai/
- ollama: https://ollama.com/
- perplexity.ai: https://www.perplexity.ai/
- poe.ai: https://poe.com/
- runpod.ai: https://www.runpod.io/
- rundiffusion.com: https://rundiffusion.com/
- runwayml.com: https://runwayml.com/
- Stable diffusion: https://stability.ai/stable-diffusion
- DeepL: https://www.deepl.com/
- you.com: https://you.com/

Future prospects

The design project is not finished. The website is designed in such a way that it can be updated dynamically. New functions and findings can be seamlessly integrated. The possibility of an autonomously updating website is not ruled out for the future.

Bibliography

AGIT. (2007). *Property Rights Manual.* AGIT. https://www.agit.de/fileadmin/Media/Für_Gruender/Schutzrechte-Handbuch/Schutzrecht_Handbuch. pdf retrieved on September 27, 2024.

Akten, M., Fiebrink, R., & Grierson, M. (2020). *Deep meditations: Controlled navigation of latent space*. arXiv. https://doi.org/10.48550/arXiv.2003.00910

Anadol, R. (2023). *AI, NFTs, and the future of art: An interview with Refik Anadol.* Boulevard. https://boulevard.co/post/interview-refik-anadol-ai-nft-art/ accessed September 27, 2024.

Association for Computational Creativity (2020). *Computational Creativity.* Association for Computational Creativity. https://computationalcreativity.net/home/about/computational-creativity/ accessed on September 27, 2024.

Atreja, D. (2024.). Thebicycleai.com. https://www.thebicycleai.com/broader retrieved September 27, 2024.

Backblaze. (2023). *AI 101: GPU vs. TPU vs. NPU.* Backblaze Blog. https://www.backblaze.com/blog/ai-101-gpu-vs-tpu-vs-npu/ Retrieved on September 27, 2024.

Barthes, R. (2007). *The death of the author*. In F. Jannidis et al. (Eds.), Texte zur Theorie der Autorschaft (pp. 185-193). Stuttgart: Reclam.

Berlich, S. (2022). *Posthumanism and pop music: Two samples: Gri- mes' We Appreciate Power and Holly Herndon's Eternal*. In S. Berlich, H. Grevenbrock, & K. Scheerer (Eds.), Where are we now? - Orientations after Postmodernism (pp. 273-285). transcript Verlag. https://doi.org/10.14361/9783839462560

Boden, M. (2010). *Can computer models help us to understand human crea- tivity?* National Humanities Center. https://nationalhumanitiescenter.org/ on-the-human/2010/05/can-computer-models-help-us-to-understand-human-creativity/ accessed September 27, 2024.

Boden, M. A. (1998). *Creativity and artificial intelligence*. Artificial Intelligence, 103(1-2), pp. 347-356. https://doi.org/10.1016/S0004-3702(98)00055-1

Boden, M. A. (2014). *Creativity and artificial intelligence: A contradiction in terms?* In E. S. Paul & S. B. Kaufman (Eds.), The philosophy of creativity: New essays (online edn). Oxford Academic. https://doi.org/10.1093/acprof:oso/9780199836963.003.0012

Botha, C. F. (2023). *Creativity and AI: A response to Boden*. In R. Hakli, P. Makela, & J. Seibt (Eds.), Social robots in social institutions - Proceedings of Robophilosophy 2022 (pp. 204-210). IOS Press BV. https://doi.org/10.3233/FAIA220637

Bown, O. , Brown, A. (2018), *Interaction design for metacreative systems*. In M. Filimowicz & V. Tzankova (Eds.), New directions in third wave human-computer interaction I, Dordrecht, Springer, pp. 67-87. https://doi.org/10.1007/978- 3-319-73356-2_5

Buehrle, A. (2008, October). *Expressionist painting and society.* La Clé des Langues. ENS de Lyon/DGESCO. https://cle.ens-lyon.fr/allemand/arts/peinture-et-sculpture/expressionist-painting-and-society Retrieved September 27, 2024.

Butler, J., Holden, K., & Lidwell, W. (2010). *Universal principles of design, revised and updated: 125 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design.* Rockport Publishers.

Callaway, E. (2024). *AI now beats humans at basic tasks - new benchmarks are needed, says major report*. Nature, 628, pp. 700-701 (2024). https://doi.org/10.1038/d41586-024-01087-4

Carey, B. (2008). *The social mind: Language and other human tricks.* Observer, 21(10). https://www.psychologicalscience.org/observer/the-social-mind-language-and-other-human-tricks retrieved on September 27, 2024.

catboxanon. (2024, March 11). *Stable Diffusion web UI: Features.* GitHub. https:// github.com/AUTOMATIC1111/stable-diffusion-webui/wiki/Features accessed September 27, 2024.

Caylor, D. (2024.). ComfyUI Nexus. [GitHub repository]. GitHub. https://github. com/daxcay/ComfyUI-Nexus Retrieved September 27, 2024.

Clackson, A. (2024, September 17). *Unlike AI, I have feelings.* I recently... [LinkedIn post]. LinkedIn. https://www.linkedin.com/posts/aelfleda_unlike-ai-i-have-feelings-i-recently-activity-7237722046195994625-FbEo Retrieved on September 27, 2024.

Crowson, K. (2022). *Simulacra aesthetic models* [GitHub repository]. GitHub. https://github.com/crowsonkb/simulacra-aesthetic-models Retrieved on September 27, 2024.

Davis, N. (2021). *Human-Computer Co-Creativity*. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 9(6), 9-12. https://doi.org/10.1609/aiide.v9i6.12603

D'Isa, F. (2023, April 16). Can algorithms make art? On Eduardo Navas's "The Rise of Metacreativity". Los Angeles Review of Books. https://www.lareviewofbooks.org/article/article-can-algorithms-make-art-on-eduardo-navass- the-rise-of-metacreativity/ accessed September 27, 2024.

Donald, M. (2020). *Arts and human symbolic cognition: Art is for social communication.* In A. Abraham (Ed.), The Cambridge handbook of the imagination (pp. 535-549). Cambridge University Press. https://doi.org/10.1017/9781108580298.036

Dunbar, R. I. M. (1992). *Co-evolution of neocortex size, group size and language in humans*. Behavioral and Brain Sciences (Unedited preprint). Cambridge University Press. Cited at https://web.archive.org/web/20090501083939/http://watarts.uwaterloo.ca/~acheyne/dunbar.html retrieved on Septem- ber 17, 2024.

Duymedjian, R. & Rüling, C.-C. (2010). *Towards a foundation of bricolage in organization and management theory*. Organization Studies, 31(2), 133-151. https://doi.org/10.1177/017084060934705

Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). *CAN: Creative adversarial networks generating art by learning about styles and deviating from style norms.* arXiv. https://doi.org/10.48550/arXiv.1706.07068

Epstein, Z., Levine, S., Rand, D. G., & Rahwan, I. (2020). *Who gets credit for AI-generated art?* iScience, 23(9), 101515. https://doi.org/10.1016/j.isci.2020.101515

Flusser, V. (2000). *Towards a philosophy of photography*. Reaktion Books.

freeCodeCamp. (2020, January 6). Brute force algorithms explained. freeCodeCamp. https://www.freecodecamp.org/news/brute-force-algorithms-explained/ retrieved on September 27, 2024.

Gabsi, A. E. H. (2024). *Integrating artificial intelligence in industry 4.0: Insights, challenges, and future prospects - a literature review.* Annals of Operations Research. https://doi.org/10.1007/s10479-024-06012-6

Gelernter, D. (1994). *The muse in the machine: Computerizing the poetry of human thought.* The Free Press.

Gesikowski, C. (2024, February 26). *Why MidJourney is so darn good.* Algography Art. Medium. https://medium.com/algography-art/why-midjourney-is-so-darn-good-eea37694c187 Retrieved September 27, 2024.

Gobble, M. M. (2019). *The road to artificial general intelligence.* Research-Technology Management, 62(3), 55-59. https://doi.org/10.1080/08956308.2019.1587336

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative adversarial nets.* Advances in Neural Information Processing Systems, 27, 2672-2680. https://doi.org/10.48550/arXiv.1406.2661

Grassegger, H., & Krogerus, M. (2023, November 4). *Bringing humanity to its full potential.* The Magazine, (44), p. 10.

Grunwald, A. (2024). *Human images and the relationship to technology and machines.* In: Zichy, M. (eds) Handbuch Menschenbilder. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-32128-4_53

Ho, J., Jain, A., & Abbeel, P. (2020). *Denoising diffusion probabilistic models.* arXiv preprint arXiv:2006.11239. https://doi.org/10.48550/arXiv.2006.11239

Horton Jr, C.B., White, M.W. & Iyengar, S.S. *Bias against AI art can enhance perceptions of human creativity*. Sci Rep 13, 19001 (2023). https://doi.org/10.1038/s41598-023-45202-3

IBM. (n.d.). *AI hallucinations.* IBM. https://www.ibm.com/topics/ai-hallucinations accessed September 27, 2024.

Illustratoren Organization e.V. (2023). *Survey on the professional situation in 2023.* https:// illustratoren-organisation.de/wp-content/uploads/2020/06/Umfrage-zur- Berufssituation-2023.pdf retrieved on September 27, 2024.

Inglewood, A. (2023, March 16). *What is a Stable Diffusion seed and how to use it.* Once Upon an Algorithm. https://onceuponanalgorithm.org/guide- what-is-a-stable-diffusion-seed-and-how-to-use-it/ retrieved on September 27, 2024.

Jenka. (2023, March 27). *AI and the American smile*. Medium. https://medium. com/@socialcreature/ai-and-the-american-smile-76d23a0fbfaf Retrieved September 27, 2024.

Jordanous, A. (2014). *What is computational creativity?* The Creativity Post. https://www.creativitypost.com/science/what_is_computational_creati- vity Retrieved September 27, 2024.

Kant, I. (1790). *Critique of the power of judgment.* Hamburg: Felix Meiner Publishers

Karger, R. (2024, January 11). *Machine intelligence and the human monopoly*. German Research Center for Artificial Intelligence (DFKI). https://www.dfki.de/web/news/maschinelle-intelligenz-und-das-mensch- liche-monopol retrieved September 27, 2024.

Kemene, E., Valkhof, B., & Tladi, T. (2024, July 22). *Generative AI's energy emis- sions: Will it help or hurt the planet?* World Economic Forum. https://www.

weforum.org/agenda/2024/07/generative-ai-energy-emissions/ retrieved on September 27, 2024.

Kingma, D. P., & Welling, M. (2013). *Auto-encoding variational Bayes.* arXiv preprint arXiv:1312.6114. https://doi.org/10.48550/arXiv.1312.6114

Kingma, D. P., & Welling, M. (2019). *An introduction to variational autoencoders.* Foundations and Trends® in Machine Learning, 12(4), pp. 307-392. https://arxiv.org/pdf/1906.02691 accessed September 27, 2024.

Lange, B. (2002). *Social sculpture*. In H. Butin (Ed.), DuMont's dictionary of terms on contemporary art (p. 276). DuMont.

Lévi-Strauss, C. (1962): *La pensée sauvage*. Plon, p. 16

Lorenz, M. (2024). *The benefits of Flexible Visual Systems.* Flexible Visual Sys- tems. https://flexiblevisualsystems.info/resources/benefits-of-fvs/ R e t r i e v e d September 27, 2024.

Manske, A., & Schnell, C. (2018). *Work and employment in the cultural and creative industries.* In F. Böhle, G. G. Voß, & G. Wachtler (Eds.), Handbook of the sociology of work (pp. 1-20). Springer. https://doi.org/10.1007/978-3-658-21704-4_14

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.* Stanford University. https://www-formal.s t a n f o r d . e d u / j m c / h i s t o r y / d a r t m o u t h / d a r t - mouth.html retrieved September 27, 2024.

McCloud, S. (2001). *Reading comics the right way.* Carlsen Verlag.

McCormack, J., & Samsel, F. (2024). Jon McCormack: *Art infused with [artificial] intelligence.* IEEE Computer Graphics and Applications, 44(2), pp. 46-54. https://doi.org/10.1109/MCG.2023.3348588

Moritz, W. (1997, August). *John Whitney: Digital harmony*. Animation World Magazine, 2(5). Cited from https://www.awn.com/mag/issue2.5/2.5pages/2.5moritzwhitney.html retrieved September 27,

2024.

Navas, E. (2012). *Remix Theory: The Aesthetics of Sampling*. SpringerViennaNewYork.

Navas, E. (2021). *Machine learning and remix: Self-training selectivity in digital art practice.* In T. Gartmann & M. Schäuble (Eds.), Studies in the arts - New perspectives on research about, in and through art and design (pp. 191- 204). De Gruyter. https://doi.org/10.1515/9783839457368-013

Navas, E. (2022). *The rise of metacreativity:* AI aesthetics after remix. Routledge.

Ochmanek, A. (2016). *Jackson Pollock*. Museum of Modern Art. https://www.moma.org/artists/4675 Retrieved September 27, 2024.

OpenAI. (2023). *GPT-4 research. OpenAI.* https://openai.com/index/gpt-4-research/ retrieved on September 27, 2024.

OpenAI. (2024). *Learning to Reason with LLMs. OpenAI.* https://openai.com/index/learning-to-reason-with-llms/ accessed September 27, 2024.

Palacios, B., Rosario, A., Wilhelmus, M. M., Zetina, S., & Zenit, R. (2019). *Pollock avoided hydrodynamic instabilities to paint with his dripping technique*. PLOS ONE, 14(10), e0223706. https://doi.org/10.1371/journal.pone.0223706

Pointon, M. (1993). *Portrayal and the Search for Identity*. Reaction Books.

Poschauko, T., & Poschauko, M. (2013). *Nea Machina: The creative machine.* Hermann Schmidt Verlag.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., & Krueger, G. (2021). *Learning transferable visual models from natural language* supervision. arXiv. https://arxiv.org/abs/2103.00020

Reckwitz, A. (2013). *The invention of creativity* (3rd ed.). Suhrkamp, cited in Reckwitz, A. (2013). The invention of creativity. Kulturpolitische Mitteilungen, (141).

Rixford, E. (2014). *Figures in the Fourth Dimension. Automata and clockwork.* http://www.figuresinthefourthdimension.com/html/Automata-clockwork.html retrieved September 27, 2024.

Royal Academy of Arts. (2018, October 10). *Demis Hassabis: Creativity and AI* - The Rothschild Foundation Lecture [Video]. YouTube. https://www.youtube. com/watch?v=d-bvsJWmqlc&t=2816s accessed September 27, 2024.

Runco, M. A. (2023). *Updating the standard definition of creativity to account for the artificial creativity of* AI. Creativity Research Journal. https://doi.org/10.1080/10400419.2023.2257977

Runco, M. A., & Jaeger, G. J. (2012). *The standard definition of creativity.* Creativity Research Journal, 24(1), pp. 92-96. https://doi.org/10.1080/10400419.2012.650092

Salk Institute. (2016). *Memory capacity of brain is 10 times more than previously thought.* Salk Institute. https://www.salk.edu/de/news-release/memory-capacity-of-brain-is-10-times-more-than-previously-thought/ accessed September 27, 2024.

Scheinberger, F. (2013). *100 ways to paint a bird.* Published by Hermann Schmidt.

Schmidt, R. F., & Lang, F. (Eds.). (2019). *Human physiology: with patho-physiology* (32nd ed.). Springer.

Siegler, R., Eisenberg, N., DeLoache, J., Saffran, J., & Pausen, S. (2021). *Developmental psychology in childhood and adolescence.* Springer Science Publishers.

Somers, M. (2019). *Emotion AI, explained.* MIT Media Lab. https://www.media.mit.edu/articles/emotion-ai-explained/ accessed September 27, 2024.

Stability AI. (2022, August 10). *Stable Diffusion public release.* Stability AI. https://stability.ai/news/stable-diffusion-announcement accessed on September 27, 2024.

Stiegler, S. (2023). *Your brain is a massively connected, ever-dynamic wonder.* Psychology Today. https://www.psychologytoday.com/us/blog/between-cultures/202301/your-brain-is-a-massively-connected-ever-dynamic-wonder accessed September 27, 2024.

Tamagnini, F. (2023). *The staggering complexity of the human brain.* Psychology Today. https://www.psychologytoday.com/us/blog/consciousness-and-beyond/202309/the-staggering-complexity-of-the-human-brain Retrieved September 27, 2024.

The AI GRID. (2024, January 12). *Sam Altman just revealed key details about GPT-5*... (GPT-5 robot, AGI + more) [Video]. YouTube. https://www.youtube. com/watch?v=JuYLCfb7CK8&t=990s Retrieved September 27, 2024.

Toffler, A. (1983). *The third wave: Opportunities for the future. Perspectives for the society of the 21st century* (transl., Die dritte Welle 1980). Goldmann.

Turing, A. M. (1950). *Computing machinery and intelligence.* Mind, 59(236), 433-460. https://doi.org/10.1093/mind/LIX.236.433

United Nations, Department of Economic and Social Affairs, Population Division. (2022). *World Population Prospects 2022: Summary of Results.* https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_zusammenfassung_der_ergebnisse.pdf retrieved on September 27, 2024.

Venners, B. (2003, March 24). *Abstraction and detail: A conversation with Andy Hunt and Dave Thomas,* Part IV. Artima. https://www.artima.com/articles/abstraction-and-detail Retrieved September 27, 2024.

Vermeulen, T., & van den Akker, R. (2010). *Notes on metamodernism.* Journal of

Aesthetics & Culture, 2(1), 5677. https://doi.org/10.3402/jac.v2i0.5677

Vincent, J. (2022, September 15). *Anyone can use this AI art generator - that's the risk.* The Verge. https://www.theverge.com/2022/9/15/23340673/ai-image-generation-stable-diffusion-explained-ethics-copyright-data accessed September 27, 2024.

West, S. (2004). *Portraiture*. Oxford University Press.

Wong, A. (2024a, June 9). *How does Stable Diffusion work?* Stable Diffusion Art. https://stable-diffusion-art.com/how-stable-diffusion-work/ accessed September 27, 2024.

Wong, A. (2024b, May 12). *Samplers in Stable Diffusion.* Stable Diffusion Art. https://stable-diffusion-art.com/samplers/ Retrieved on September 27, 2024.

Wong, A. (2024c, January 23). *Denoising strength in Stable Diffusion*. Stable Diffusion Art. https://stable-diffusion-art.com/denoising-strength/ Retrieved September 27, 2024.

Wong, A. (2024d, July 7). *ControlNet in Stable Diffusion*. Stable Diffusion Art. https://stable-diffusion-art.com/controlnet/ Retrieved on September 27, 2024.

Wong, A. (2024e, September 15). *How to train LoRA in Stable Diffusion*. Stable Diffusion Art. https://stable-diffusion-art.com/train-lora/ Retrieved on September 27, 2024.

Woodall, J. (Ed.). (1997). *Portraiture: Facing the Subject.* Manchester University Press.

Zhang, L., Wu, X., Zhang, Y., Zhu, Y., Wu, Z., & Wang, Y. (2023). *ControlNet: Adding Conditional Control to Text-to-Image Diffusion Models.* arXiv preprint arXiv:2302.05543. https://doi.org/10.48550/arXiv.2302.05543

List of illustrations

Illustrations not specifically mentioned in the text are by the author. Fig. 1

Illustrators Organization e.V. (2023). *Survey on the professional situation in 2023.* https://illustratoren-organisation.de/wp-content/uploads/2020/06/ Umfrage-zur-Berufssituation-2023.pdf, p. 9 retrieved on September 27, 2024.

Fig.2

OpenAI. (2024). *Learning to Reason with LLMs*. OpenAI. https://openai.com/ index/learning-to-reason-with-llms/ accessed September 27, 2024.

Fig. 3

Marsyas. (2007). *Antikythera Mechanism.* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:NAMA_ Machine_e%27Anticyth%C3%A8re_1.jpg (Creative Commons Attribution-ShareAlike 4.0 International) retrieved on September 27, 2024.

Fig. 4

Daderot. (2011). *Henri Maillardet automaton, London, England, c. 1810 - Franklin Institute.* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Henri_Maillardet_automaton,_London,_England,_c._1810_-_Franklin_Institute_-_DSC06656.jpg (Creative Commons Attribution-ShareAlike 4.0 International) retrieved on September 27, 2024.

Fig. 5

Takahshi, Y., & Okuno, H. G. (2006). *Design and Implementation of Programmable Drawing Automata*. In Proceedings of the ICRA. Image source: https://www.semanticscholar.org/paper/Design-and-Implementation-of-Programmable-Drawing-Takahashi-Okuno/328d4c5fba3cb99f5583385248dc031d5e4549fd retrieved on September 27, 2024.

Fig. 6

Kingma, D. P., & Welling, M. (2019). *An introduction to variational autoencoders.* Foundations and Trends® in Machine Learning, 12(4), p. 58. https://arxiv.org/pdf/1906.02691 accessed September 27, 2024.

Fig. 7

Syned. (2019). *MNIST: Reborn, Restored, and Expanded.* https://syncedreview.com/2019/06/19/mnist-reborn-restored-and-expanded-additional-50k-training-samples/ retrieved on September 27, 2024.

Fig. 8

Zhang, M. M., Gundersen, G. W., & Engelhardt, B. E. (2024). *Bayesian Non-linear Latent Variable Modeling via Random Fourier Features.* ResearchGate. https://www.researchgate.net/publication/37160

5629_Bayesian_Non-linear_Latent_Variable_Modeling_via_Random_Fourier_Features/

figures?lo=1 (Creative Commons Attribution-ShareAlike 4.0 International) retrieved on September 27, 2024.

Fig. 9
Kingma, D. P., & Welling, M. (2013). *Auto-encoding variational Bayes.* arXiv preprint arXiv:1312.6114. https://doi.org/10.48550/arXiv.1312.6114, p. 10

Fig. 10
Pamputt. (2018). *Portrait of Edmond de Belamy.* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Edmond_de_Belamy.png (Creative Commons Attribution-ShareAlike 4.0 International) retrieved on September 27, 2024.

Fig. 11
Ras67. (2017). *Duchamp Fountain.* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Duchamp_Fountaine.jpg (Creative Commons Attribution-ShareAlike 4.0 International) retrieved September 27, 2024.

Fig. 12
Arkesteijn, J. (2015). *Keith Haring (1986).* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Keith_Haring_(1986).jpg (Creative Commons Attribution-ShareAlike 4.0 International) retrieved September 27, 2024.

Fig. 13
XRay. (2018). *Elbphilharmonie in Hamburg, HafenCity.* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Hamburg,_HafenCity,_Elbphilharmonie_--_2016_--_3062_(bw).jpg (Creative Commons Attribution-ShareAlike 4.0 International) retrieved on September 27, 2024.

Fig. 14
Vysotsky. (2023). *Boris Eldagsen.* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:EldagsenElectrician.jpg (Creative Commons Attribution-ShareAlike 4.0 International) retrieved September 27, 2024.

Fig. 15
File Upload Bot. (2005). *Narcissus by Caravaggio (1594-96).* Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Michelangelo_Caravaggio_065.jpg (Creative Commons Attribution-ShareAlike 4.0 International) retrieved September 27, 2024.

Fig. 16
Saunders, B. L. (2008). *Self-portrait drawings under the influence of drugs.* Bryan Lewis Saunders. http://bryanlewissaunders.org/sps/ accessed on September 27, 2024.

Fig. 17
Unlimited Dream Co (2024). *Writing good VQGAN+CLIP prompts part two - artist and genre modifiers.* Unlimited Dream Co. https://www.unlimiteddreamco.xyz/articles/writing-good-prompts-part-2 accessed September 27, 2024.

Fig. 18
Refsgaard, A. (2020). *Erasing & Enhancing Essentials.* https://www.andreasrefsgaard.dk/projects/erasing-enhancing-essentials/ Retrieved on September 27, 2024.

Fig. 19
Anadol, R. (2022). *Unsupervised: Machine Hallucinations, MoMA.* Refik Anadol Studio. https://refikanadolstudio.com/projects/unsupervised-machine- hallucinations-moma/ retrieved September 27, 2024.

Fig. 20
Fernandez, J. (2024). *Jumoke Studio Projects.* https://jumoke.studio/ Retrieved September 27, 2024.

Fig. 21
Hellandsjø, K. (n.d.). *Jackson Pollock biography.* Store Norske Leksikon. https://snl.no/Jackson_Pollock (Creative Commons Attribution-ShareAlike 4.0 International) retrieved September 27, 2024.

Fig. 22
AGIT. (2007). *Property Rights Manual.* AGIT. https://www.agit.de/fileadmin/Media/Für_Gruender/Schutzrechte-Handbuch/Schutzrecht_Handbuch. pdf, p. 9 retrieved on September 27, 2024.

I hereby declare that I have written this thesis independently and without the use of any aids other than those specified. I have indicated the source of any ideas taken verbatim or in spirit from other works. This also applies to illustrations and sources from the Internet.

I agree to the inspection and borrowing of this document in the library of the HAW Hamburg, Department of Design.

Place, date, signature

Hamburg, September 27, 2024