SYNTHESIZING MEANING

DESIGNING A PRODUCTION PIPELINE FOR THE ARTISTIC APPROPRIATION OF GENERATIVE MACHINE LEARNING TO MEDIATE SPECULATIVE CONCEPTS

SYNTHESIZING MEANING

DESIGNING A PRODUCTION PIPELINE FOR THE ARTISTIC APPROPRIATION OF GENERATIVE MACHINE LEARNING TO MEDIATE SPECULATIVE CONCEPTS

FINAL THESIS
MA OF INTEGRATED DESIGN
KÖLN INTERNATIONAL SCHOOL OF DESIGN,
TH KÖLN, FACULTY OF CULTURAL STUDIES

PUBLISHED
31TH JANUARY, 2022
BY LISA MARLEEN MANTEL
AND LAURA WAGNER

SUPERVISED BY
PROF. NINA JURIC
PROF. DR. LAURA POPPLOW
PROF. DR. LASSE SCHERFFIG

4
PROLOGUE

7
ABSTRACT

8
INTRODUCTION

10
POSTHUMANSIM VS. TRANSHUMANISM
FUTURE OF HUMANITY
QUESTIONING THE CARTE-SIAN WAY OF THINKING
EMBODIED COGNITION

18
WORLD TO LATENT REPRESENTATION
SYMBOLIC AI, INTERNET AI, EMBODIED AI
EXTRAPOLATED THINKING
PROPAGATING BIASES

30
DESIGN PROFESSION IN FLUX
MACHINE CREATOR – HUMAN CURATOR
SPARKS OF IMAGINING
APPLIED HUMAN-MACHINE COLLABORATION

40
MEDIATING A SPECULATIVE CONCEPT
MOVING BEYOND THE HUMAN/ MACHINE BINARY
TRANSMACHINIST APPROACH

46
BORROWED LIMBS – A SHORTFILM

48
CREATING A FILM WITH AN UNRULY ACTOR
LET’S NOT SERVE THE TOPOS
A LATENT SPACE FULL OF EMOTIONS
THE TOOLBOX
BEING IN A LOOP OF DIALOGUE
DESIGNING A PIPELINE FOR MACHINE COLLABORATION

100
CONCLUSION

104
EPILOGUE

106
BIBLIOGRAPHY

112
TABLE OF ILLUSTRATIONS

116
APPENDICES

HUMAN VS. ***HUMAN***

In accordance with posthumanist thinkers such as Rosi Braidotti, Donna Haraway, and Katherine Hayles, we distinguish how we use the terms human and ***human*** (written in italics). In a humanist societal context, ***human*** is not to be understood as a neutral term.[01] It has historically been used to refer exclusively to white, hetero, cis males and excludes so-called missing people (among others women, Black, First Nation, dark- and darker, African, indigenous, LGBTQ+, neurodiverse and people with disabilities). However, in a biological and evolutionary context, we refer to the species homo sapiens sapiens as human.

01 – Braidotti, Rosi, Hlavajova, Maria and contributors, „Posthuman Glossary," Posthuman Critical Theory, London: Bloomsbury Academic, 2018

FIG 01 – *Human*
StyleGAN2-ada-pytorch

FIG 02 – Missing People
StyleGAN2-ada-pytorch

Thesis Keywords:
post- and parahuman, embodiment, decentering the humanness being, designerly appropriation of ai tools, machine creator – human curator, generative machine learning, world to latent space, synthetic images

ABSTRACT

In times of ubiquitous computing, artificial intelligence, and virtual ***humans***, it seems like an obvious next step for techno-optimists to see their chance for immortality in technological advancements. These future narratives are almost exclusively based on cartesian notions of disembodied intelligence. Influenced by posthumanist theorists, transhumanist approaches that envision the ***human*** body enhanced and augmented with machine parts are not only questioned but appropriated and inverted with the use of speculative design. This work aims to emphasize the mediating role speculative design can have in the face of disruptive technological advancements, as it allows for the exploration and simulation of ethical issues and provokes discussion about current developments and how they will influence our future way of life.

The theoretical work results in a short film created with a designerly appropriation of AI tools. The suggested production pipeline, developed in the process, uses a myriad of AI tools. The main part of the footage was produced with a novel machine-learning-backed process called CLIP-guided diffusion that allows for image synthetization based on language prompts that are carefully refined in a human-machine loop of communication: prompting the machine, evaluating the result, and refining the next prompt in order to get the anticipated result. In a way, the practical processes involved in making the film reflect on the speculative ideas conveyed in it: the protagonist is an AI that utilizes a human body to gain an embodied understanding of its environment, a result of consequent machine-human hybridization - a posthuman being.

INTRODUCTION

The ability to simulate new concepts will become increasingly crucial to the future role of designers. When technological progress is accelerating and the possibilities of appropriation are given, it is even more important to point out the possible risks of future developments, preferably by the same means. Potential risks discussed in the first part of this thesis are derived from posthumanist thinkers who criticize a techno-optimism that does not deal with the consequences of its methods. Visions of a detachment of the mortal body completely disregard that thinking is dependent on an organic body. Interactions of a social and material kind, necessary to form a mental model of the environment and oneself, are ignored. In posthumanist opinions, a human body diversified by technology and hybridized with objects, animals, or nature can lead to a new form of subjectivity.

The second part will discuss how a part of the researchers and developers involved in AI technologies follow the dictum that AI can overcome all its limitations if fed with enough raw data; no or little effort is needed to make the AI system more sophisticated. Data omission and simplification for the sake of computational feasibility introduce and amplify bias. Once a model is trained with biased data, it serves as a multiplier. Models trained today will inadvertently determine what future content will look like, as they continuously add new data to our collective internet knowledge, which will be used to train new representations.

The third focus of this work will be how design, motivated by posthumanist visions, can be a catalyst to spark further thinking. Rather than focusing on rapid, market-driven implementation, speculative design decouples from industrial agendas. Thought experiments visualized and narrated with design methods can simulate future developments. The emergence of new technologies will also be analyzed regarding the design profession and possible integrations with the skill-and mindset of designers.

A speculative design concept will be developed and visualized in a short film derived from the theoretical research. Notions will be reflected in a collaborative design method that will appropriate and explore a myriad of AI tools, all of them released to the public less than a year ago.

The last part proposes a modular design pipeline for controlled image synthesis with AI tools as designerly appropriation to mediate speculative concepts. The best practice outcome for image synthetization, the short film, will be produced with a novel machine-learning-backed design method. This selected design method is based on iteration and collaboration with the machine as a new collaborator of a different species. While appropriating and exploring these tools, a critical reflection on their nature, purpose, and possible impact on society will be upheld.

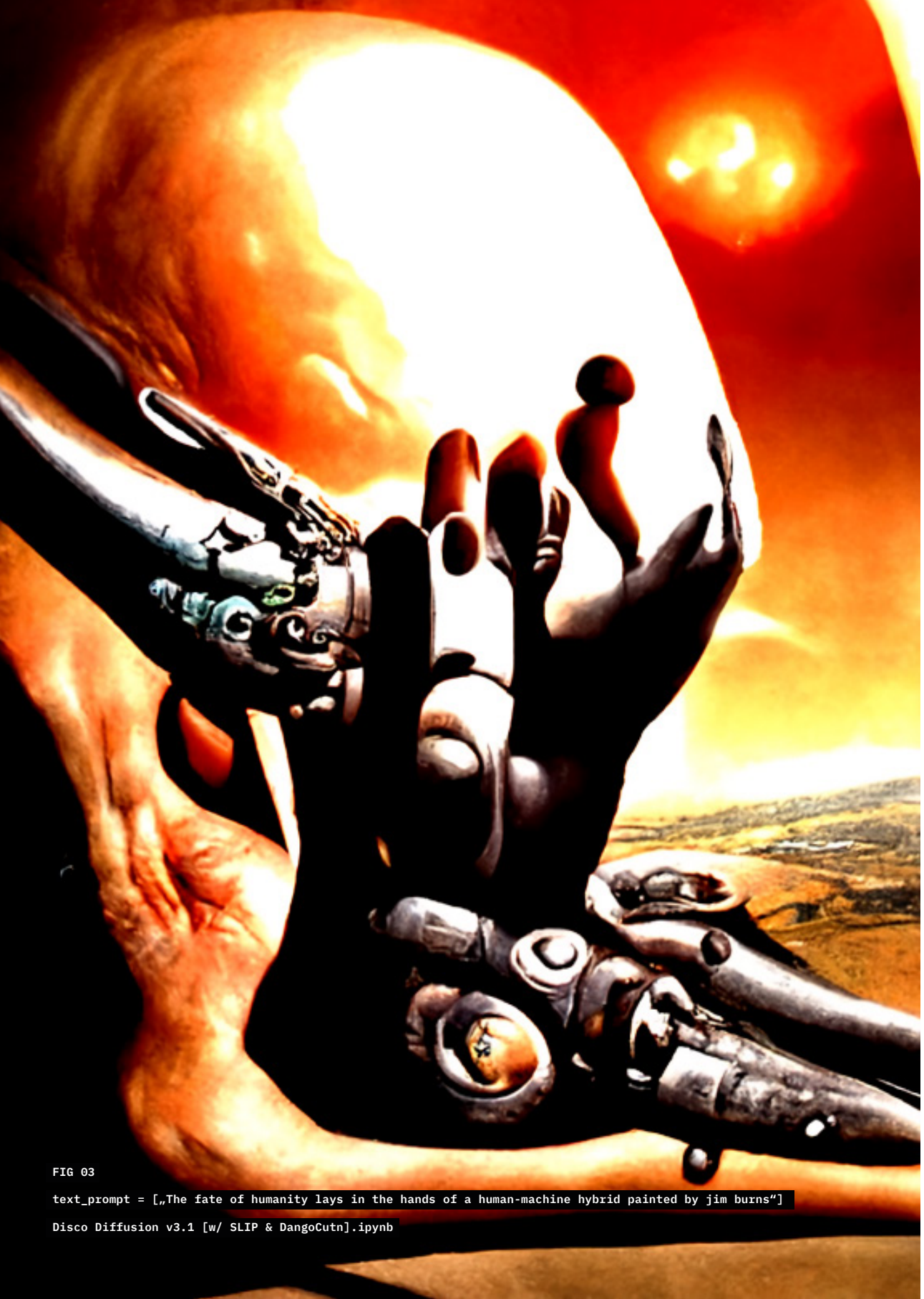


FIG 03

text_prompt = [„The fate of humanity lays in the hands of a human-machine hybrid painted by jim burns“]

Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

POSTHUMANISM VS. TRANSHUMANISM

There is an inherent urge in ***human*** beings to feel special. Although the cosmic perspective undoes this urge, the depictions of being the coronation of creation echo in humanist societies. Unfortunately, they seem not entirely flawless: the human body is transience. In religious contexts, this flaw is weakened by the promise of paradise after death. However, in times of ubiquitous computing, artificial intelligence, and virtual ***humans***, it seems like an obvious next step for techno-optimists to see their chance for immortality in technological advancements.

FUTURE OF HUMANITY

The Cryonics Institute[02] created a business model for wealthy ***humans*** freezing their bodies. In tanks filled with liquid nitrogen, they are stored until medical technology can revive them in an unknown future. The Moravec Mind Transfer offers a complete detachment of the mortal body.[03] Hans Peter Moravec is an Austrian futurist who worked in the field of robotics and artificial intelligence. In his book Mind Children (1988), he suggests transforming the ***human*** mind into a mechanical structure of nanobots. The nanobots are copies of the neurons of a human brain simulating it at the end of the process entirely, uploaded on a computer. Such transhuman approaches aim to enhance the ***human*** and secure the hegemony. Driven by dualistic beliefs[04], those concepts mainly focus on enhancing parts of the ***human*** body mechanically or erasing the body entirely and only preserving the brain. Transhumanism as a no-

02 – "Cryonics Institute," accessed January 09, 2022, https://www.cryonics.org/.

03 – N. Katherine Hayles, How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics (Chicago, Ill: University of Chicago Press, 1999), 1.

04 – Ontological dualism states that body and mind are two forms of existence that are incompatible with each other and exist separately.

tion is the fulfillment of the enlightenment project, which was mainly the perfectibility of the ***human*** through the application of scientific reason situated in the seventeenth century in Europe. One of the most famous representatives is philosopher and mathematician René Descartes, who hierarchized the mind and body, assuming that only humans can think. And not even among the human species, everybody is considered ***human*** by Cartesian male theorists. This designation is reserved for heterosexual, white cis males. Not included are those who contemporary philosopher and feminist Rosi Braidotti calls the missing people[05]: among other women, Black, First Nation, dark- and darker, African, indigenous, LGBTQ+, neurodiverse, and people with disabilities. Moreover, the same people are consequently forgotten to transform into a transhumanist society. In the words of Braidotti, Transhumanism is nothing else than Neo-Humanism.[06]

The Mind over Body hierarchy is also a widespread belief and argument while developing future digital tools. The mainly male-led digital field has a massive influence on technological advancements, making technology, not at all neutral. Cartesian theories are extended to the field of information technology. The transhumanist views create a basis that allows ***human*** beings to be seamlessly transferred into intelligent machines. There is no difference between: "bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology, and ***human*** goals".[07] In this modern dichotomy, white, European men are disembodied minds, whereas others are overexpressed or suppressed.[08] Dualism is used as an instrument of governance. Missing people who cannot participate in the "***human*** enhancement" will be erased just like the body of the enhanced ones.

05 – Rosi Bra idotti and Maria Hlavajova and contributors, „Posthuman Glossary," Posthuman Critical Theory, London: Bloomsbury Academic 2018, 342.

06 – Rosi Braidotti, "A Theoretical Framework for the Critical Posthumanities," Theory, Culture & Society 36, no. 6 (November 1, 2019): 32, https://doi.org/10.1177/0263276418771486.

07 – N. Katherine Hayles, „How We Became Posthuman", 3.

08 – Rosi Braidotti, The interview with Judith Butler (S. W. Oh, Trans. Korea). Journal of Feminist Theories & Practices (1999), 1, 278.

We are already augmenting, organizing, and externalizing our cognitive processes with machines. We store and query our knowledge and memories in ones and zeros and extend our perception with sensory machines that tell us in human-readable form meteorological forecast. We repair ourselves with artificial hearing aids and pacemakers. Human societies will be hyper-connected in the future, and the organic will become even more digitized. Conclusively ***humans*** could repeat oppression, exploitation, and inequality already seen in history by becoming non-biological thinking beings. Nevertheless, the exponential acceleration of the digital world is amplifying these significantly.[09]

QUESTIONING THE CARTESIAN WAY OF THINKING

A school of thought that sees itself as an abolishment of the body-mind dualism is posthumanism. It understands itself as a chance for those not seen as equally ***human*** and aims to overcome sexual, racial, and ethnic inequalities retrieved by anthropocentric history and perpetuated by transhumanist supporters. Instead of adapting, the individuals neglected by humanist structures should form transspecies alliances and companionships.[10] Yeon Jeong Gu, Soongsil University in Seoul, suggests reconfiguring a technologically mediated body as a new form of subjectivity.[11] Influenced by posthuman thinkers like Rosi Braidotti and Katherine Hayles, her vision is a human body diversified and hybridized with objects, animals, or nature through technology. In this context, hybridity can also be understood as an intersection with other organisms. An im-

09 – Yeon Jeong Gu, "The Disembodiment of Digital Subjects and the Disappearance of Women in the Representations of Cyborg, Artificial Intelligence, and Posthuman," Asian Women 36, no. 4 (December 31, 2020): 33, https://doi.org/10.14431/aw.2020.12.36.4.23.

10 – Nicholas Gane, "When We Have Never Been Human, What Is to Be Done?: Interview with Donna Haraway," Theory, Culture & Society 23, no. 7–8 (December 1, 2006): 135–58, https://doi.org/10.1177/0263276406069228.

11 – Yeon Jeong Gu, "The Disembodiment of Digital Subjects and the Disappearance of Women in the Representations of Cyborg, Artificial Intelligence, and Posthuman," 33.

portant aspect is coexistence and symbiosis with other living beings.
All this contrasts with an understanding of the world which puts one species in the center. Donna Haraway and her figures of thought like cyborgs and chimeras create a blueprint for this posthuman body. With her theoretical approaches of hybrids of machines and organisms[12], she declares that humans already fused with technology when using objects as tools. For philosopher Bernard Stiegler, objects are indispensable for transporting human knowledge. They outsource thoughts from our bodies into non-human objects.[13] Through written language or art, experiences are passed on to the next generations. By using and integrating these tools into the human body, people reshape themselves and transform. Humanity and its capabilities multiply as media becomes increasingly complex. People have a reason because they use writing and art, not the other way around.
The cultural scientist Zoë Sofoulis, on the other hand, criticizes the terms "posthuman" and "non-human objects" because they obscure the human labor invested in the creation of hybrid artifacts. The prefix "post", in her opinion, is meant to evoke newness. She proposes the term "parahuman" instead to signify that action and intelligence exist just as much outside and alongside the human. Humanity is not redundant.[14] This thesis continues to use the term "posthuman" but recognizes this criticism.
The hybridity of organisms and objects is a preferable framework to consider transformation. The human body is in constant flux, interacting with other beings and objects. This process forms and reconstructs it constantly. Thousands of years of human history affected human behaviors and thoughts.[15] From a socio-technical point of view, human personalities are often defined by connecting with technical or non-human entities. Deities or influential personalities are usually not shown alone or only naked in ancient or medieval depictions but in connection with objects such as a sword or book.[16] On the contrary of technology-centered discourses where materiality must be transcended, technology is a part of an organism's body and shapes it.
Posthuman identities require an exchange with each other but furthermore with non-human actors. Human thinking is related to the human body. This fact makes it crucial to practice interrelationships and value a symbiotic coexistence. Yeon Jeong Gu suggests experimenting with "technically mediated variant subjectivity beyond the categories of normality."[17] Katherine Hayles believes that future development should be conducted with care. Information like that of humanity cannot exist detached from a body. Embodiment gives context and meaning to information while it can be destroyed but not restored. The material world is fragile and cannot be replaced.[18] Posthumanism signals not the end of humanity but the end of a specific concept of the ***human*** being. It was a concept that at best applied to that part of humanity that had the wealth and power to conceive of themselves as autonomous beings.[19] Social and cultural structures need to be rethought with the ***humanness*** being decentered and focus on methods of communication and living through the other where boundaries are increasingly blurred.

12 – Nicholas Gane, "When We Have Never Been Human, What Is to Be Done?: Interview with Donna Haraway,"

13 – Patrick Crogan, "Bernard Stiegler: Philosophy, Technics, and Activism," Cultural Politics 6, no. 2 (July 1, 2010): 133–56, https://doi.org/10.2752/175174310X12672016548162.

14 – Zoë Souflis, "Post-, Nicht- und Parahuman", in „Future Bodies. Zur Visualisierung von Körpern in Science und Fiction", ed. Marie-Luise Angerer, Kathrin Peters und Zoë Sofoulis. 273–300, Wien, New York: Springer-Verlag 2002, 283-284.

15 – N. Katherine Hayles, „How We Became Posthuman", 284.

16 – Zoë Souflis, "Post-, Nicht- und Parahuman", 295-296.

17 – Yeon Jeong Gu, "The Disembodiment of Digital Subjects and the Disappearance of Women in the Representations of Cyborg, Artificial Intelligence, and Posthuman," 42.

18 – N. Katherine Hayles, „How We Became Posthuman", 49.

19 – see ibid. 286.

EMBODIED COGNITION

The critiques of dualism tie in seamlessly with more recent developments in the cognitive sciences. Here it is assumed that there is an interaction between cognition, sensory, and motor functions reflected in the representation of thought processes.[20] The thinking process is multimodal, which means different sensory systems work together during an interaction. Human individuals measure their movement space in units of their bodies. This way, perceptions can be translated into direct, physical action. One typical example is children learning to count. In the beginning, they use their fingers to represent numbers as an embodied translation of thoughts. Thus, the central approach of the embodiment is the emergence of consciousness through a body whose nature, functions, and embeddedness in an environment are essential to physical experiencing.[21]

Neuroscientist Antonio Damasio also links actions to emotions that we learn in socialization. He locates three abilities in the brain's frontal lobe in his hypothesis. One of them is body perception. Depending on how different interactions are experienced, we store them as pleasant or unpleasant. Damasio calls this categorization "somatic markers". They form our images, influence later decisions, and trigger subconscious preliminary decisions to warn us, for example, about things we have already had bad experiences[22] with. Through these experiences, the human brain tries to predict the near future and prepare for appropriate actions. Predictions about sensory input are being made and compared to the actual input. That way, the mental model is revised. This method is, in a way, a surprise minimization. It is called free energy minimization in cognitive science, whereas free energy is the long-term average of prediction error.[23]

This very concept of minimizing surprise is discussed with the "dark room problem". It states that a perfect agent will reduce free energy as far as possible. However, this would mean that a place has to be visited that offers little stimuli: the darkroom. Individuals do not try to avoid stimuli at all, not only because such an environment implies being without food and social interaction and thus would lead to certain death, by starvation or boredom. Organisms seem to search for stimuli instead. This stimulus seeking proves for the postdoctoral researcher Andrew Sims either that the concept of surprise minimization can be disproved or that not all stimuli are necessarily surprising.[24]

20 – Werner Stangl (2022). Stichwort: ‚Embodied Cognition – Online Lexikon für Psychologie und Pädagogik'. Online Lexikon für Psychologie und Pädagogik. https://lexikon.stangl.eu/14550/embodied-cognition, Accessed 11.01.2022

21 – Werner Stangl (2022). Stichwort: ‚Embodiment – Online Lexikon für Psychologie und Pädagogik'. Online Lexikon für Psychologie und Pädagogik. https://lexikon.stangl.eu/2175/embodiment/, Accessed 11.01.2022

22 – Werner Stangl (2022). Die Hypothese der somatischen Marker. [werner stangl]s arbeitsblätter. https://arbeitsblaetter.stangl-taller.at/GEHIRN/GehirnSomatischeMarker.shtml, Accessed 11.01.2022

23 – Andrew Sims, "The Problems with Prediction: The Dark Room Problem and the Scope Dispute" Philosophy and Predictive Processing, 2017, 4. https://doi.org/10.15502/9783958573246, Accessed 11.01.2022

24 – Andrew Sims, "The Problems with Prediction: The Dark Room Problem and the Scope Dispute", 16.

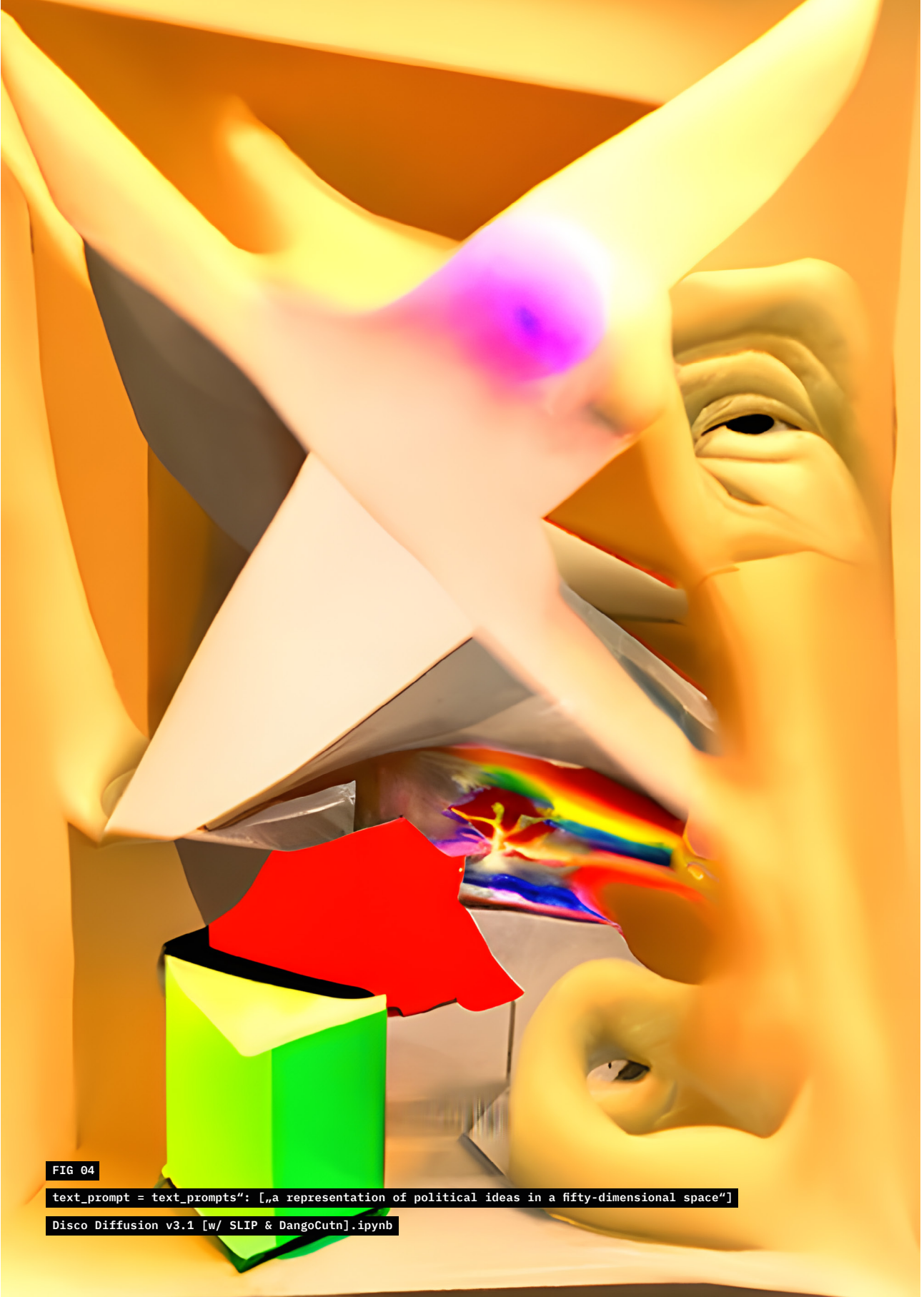


FIG 04

text_prompt = text_prompts“: [„a representation of political ideas in a fifty-dimensional space“]

Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

WORLD TO LATENT REPRESENTATIONS

In his 1951 and 1950 papers „Intelligent Machinery – a Heretical Theory“[25] and „Computing Machinery and Intelligence“[26], Alan Turing envisioned a „child machine“ that „learns by experience“ through a concept of pleasure and pain administered in an „unemotional language“ ideally by a „highly competent schoolmaster“ with whom it communicates back and forth by typewriter. He also introduces some principle of initial randomness, as his machine „will make mistakes [and] sometimes come to unexpected and interesting conclusions“. The machine will store its learnings in „indexes of experiences“. Turing addresses the possibility of embodiment but dismisses it: His child machine is disembodied; it will have no legs and cannot be ordered to go out and „fill the coal scuttle“. Even though his notions in retrospect remind us about the principles of today's learning machines, our collective internet knowledge[27] pinpoints the start of the research field that we refer to as AI six years later.

SYMBOLIC AI, INTERNET AI AND EMBODIMENT

In 1956 Marvin Minsky initiated a small workshop financed by the Rockefeller Foundation together with the by then fresh mathematics post-graduate McCarthy, who invented the term artificial intelligence as it is used today.[28] It was the first time that a group worked to create machines that can perform tasks that we would classify as requiring intelligent behavior or even (depending on the exact definitions of the problems at hand) to reverse engineer animal intelligence altogether. They used an

25 – A. M. TURING, Intelligent Machinery, A Heretical Theory*, Philosophia Mathematica, Volume 4, Issue 3, September 1996, Pages 256–260, https://doi.org/10.1093/philmat/4.3.256

26 – A. M. Turing, “I.—©MPUTING MACHINERY AND INTELLIGENCE,” Mind LIX, no. 236 (October 1, 1950): 433–60, https://doi.org/10.1093/mind/LIX.236.433.

27 – “Start of AI - Google Search,” accessed January 22, 2022, https://www.google.com/search?q=start+of+ai&oq=start+of+ai&aqs=chrome..69i57.3662j0j4&sourceid=chrome&ie=UTF-8.

28 – “Melanie Mitchell - Artificial Intelligence_ A Guide for Thinking Humans Chapter: The Roots of Artificial intelligence,Reprint, e-book, Picador, 2020

IBM 701 Electronic Data Processing Machine[29] - an electrical computer - unlike Turing, that constructed a computer entirely made from mechanical parts. That might also be why the term Artificial Intelligence prevailed the term Intelligent machines by Turing, as it was better befitting the reality of the research field today. The field of embodied cognitive science, renouncing the cartesian dualist and disembodied approach, had a relatively short history until it became well established in cognitive science and philosophy. In terms of embodiment, AI research and cognitive science seem to have gone through a similar chronology of ideas and their rejections before converging over time: The field of AI started out (1956 with the said workshop at Dartmouth college) with what today we call the symbolic approach or GOFAI - good old-fashioned AI as the dominant approach to the field before it was mainly abolished and overtaken by a new approach: sub-symbolic AI. Historically the (older) symbolic AI approach formalized human-readable words and phrases and their relationships into tokens that can be rearranged by mathematical and logical operations to go from a current state to the desired state. One reason for this approach being dominant initially may have been that it was in line with the cartesian claim that we are radically distinct from and superior to animals, having a soul and uniquely possessed of the ability for abstract reason. Symbolic AI also harmonizes with the main cognitivist belief that cognition is the mere rule-based manipulation of abstract representations. After a period of AI winter, in 2012, Neural Networks made headlines with a Neural Network that won the ImageNet Visual Recognition Challenge. The research in AI started to shift to an approach that emphasizes on perception. The sub-symbolic approach, which became the dominant approach up until this day, is inspired by how neural processing in the brain works. Here, the *rules* are weights of artificial neurons that are gradually adjusted in a process called Machine Learning. Machine Learning is the process of extracting rules from a large amount of information to infer (reconstruct and generate) data or actuator-activation based on these rules. Said large amount of data is usually taken from the internet. Hence the current state of Machine Learning models is often dubbed "internet AI". While internet AI generally performs very well in specific narrow-domain tasks like object recognition or natural language processing, it is far from the ridiculously optimistic forecasts the first proponents of AI made about the technology they were working on and the prospect of singularity. Marvin Minsky proclaimed, "within a generation [...] the problems of creating artificial intelligence will be substantially solved."[30] Today's AI research field is on par with the learnings of cognitive science, which maintains evolutionary continuity and generally accepts that there is much more to cognition than mental representation. Embodied cognition, in particular, emphasizes intelligence as something that executively emerges in the interaction of an agent with an environment and as a result of homeostatic/allostatic self-regulation and sensorimotor activity.[31] Embodiment in AI means giving a tangible or visible form to a notion in contrast to the amodal representation of concepts. One foundational concept on the road to embodied AI is Continual Learning. Built on the idea of learning continuously and adaptively about the external world and enabling the autonomous incremental development of more complex skills and knowledge. In the context of Machine Learning, it means smoothly updating the prediction model to take into account different tasks and data distributions but still re-use and retain useful knowledge and skills during time. In terms of a Neural Network, that would entail a continuous changing of neuron weights to react adaptively to an ever-changing environment.

29 – "IBM 701," in Wikipedia, December 1, 2021, https://en.wikipedia.org/w/index.php?title=IBM_701&oldid=1058120081.

30 – Melanie Mitchell, „Artificial Intelligence_ A Guide for Thinking Humans", „The Roots of Artificial intelligence", Reprint, e-book, Picador, 2020 (quote taken from Autor)

31 – Shapiro, Lawrence and Shannon Spaulding, „Embodied Cognition", The Stanford Encyclopedia of Philosophy (Winter 2021 Edition), Edward N. Zalta (ed.) https://plato.stanford.edu/archives/win2021/entries/embodied-cognition

There still is relatively little consensus regarding what exactly constitutes *embodiment*. Computer Vision researchers define Embodied AI as artificial agents operating in 3D environments that base their decisions on egocentric perceptual inputs that change with agent actions. This research sub-field could be called soft embodied AI. Soft embodied AI enables the training of embodied AI agents (virtual robots and egocentric assistants) in a realistic 3D simulator before transferring the learned skills to reality. AI Habitat[32] is a research platform by the company Meta (formerly Facebook) that includes a physics-enabled, photorealistic 3D simulation environment facilitating the simulation of situated AI agents. Example use cases involve mobile agents moving through an apartment to search and retrieve objects or information. **(FIG 05)**

Of course, some approaches reject the necessity of embodiment altogether. Dataism[33] is the belief that all-natural behavior can be algorithmically inferred and replicated given that there is sufficient data. The scaling hypothesis is the dataist belief that the more data goes into training a Neural Network, the more sophisticated behavior it will display. GPT-4, the successor of GPT-3, which is commercially used to create articles, poetry, stories, news reports, and dialogue, will have 100 trillion parameters — 500x the Size of GPT-3.[34] The researchers and developers involved follow the dictum that AI can overcome all its limitations if it is fed with a sufficient amount of raw data; no or little effort is needed in making the AI system more sophisticated.

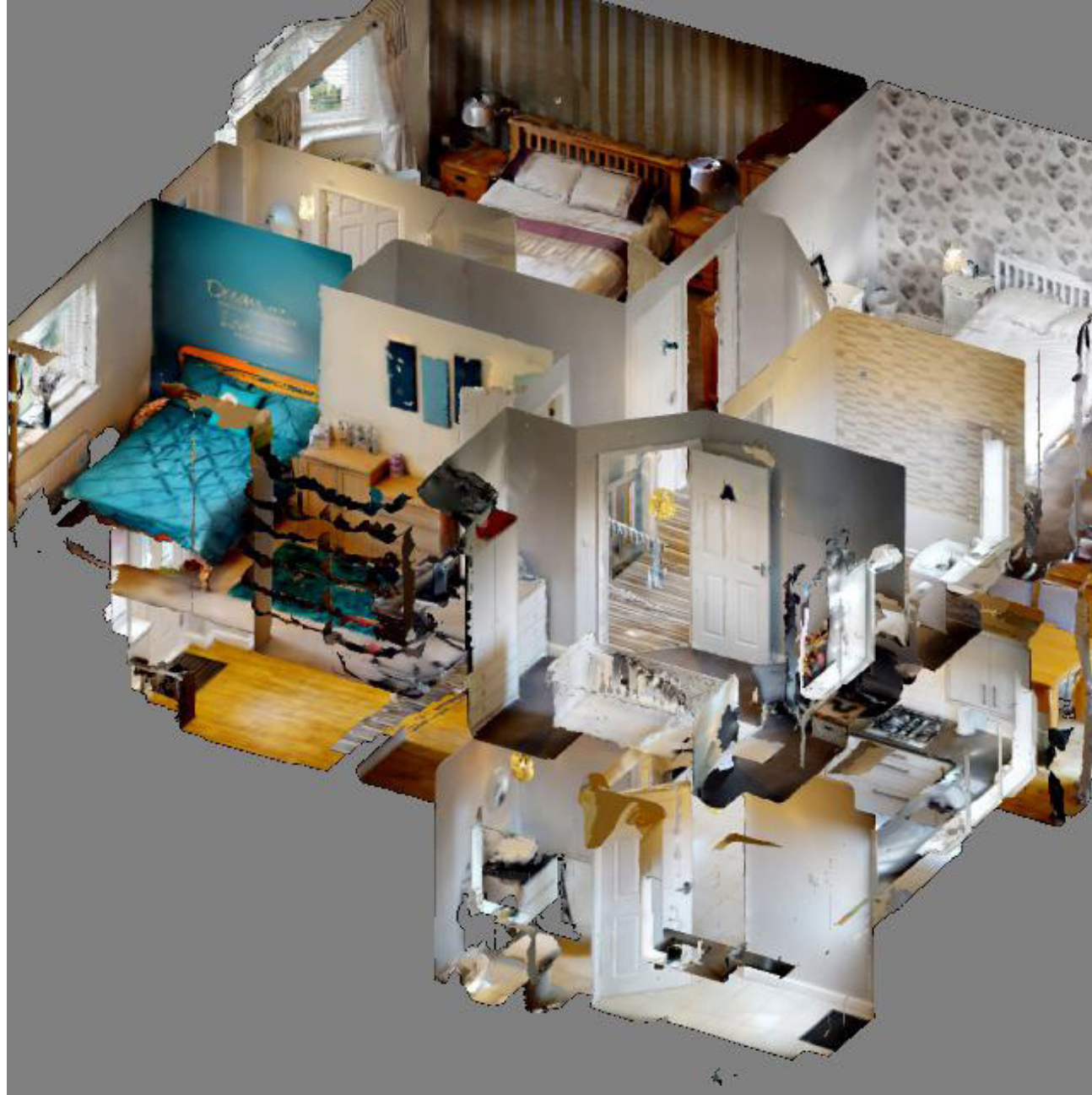


FIG 05 – AI Habitat

32 – Manolis Savva et al., "Habitat: A Platform for Embodied AI Research," in 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South): IEEE, 2019), 9338–46, https://doi.org/10.1109/ICCV.2019.00943.

33 – "Dataism," in Wikipedia, January 10, 2022, https://en.wikipedia.org/w/index.php?title=Dataism&oldid=1064746251.

34 – Alberto Romero, "GPT-4 Will Have 100 Trillion Parameters — 500x the Size of GPT-3," Medium, September 11, 2021, https://towardsdatascience.com/gpt-4-will-have-100-trillion-parameters-500x-the-size-of-gpt-3-582b98d82253.

EXTRAPOLATED THINKING

The advent of large-scale Machine Learning models might be seen in retrospect as an inflection point for the introduction of synthesized representations of cultural knowledge and imagery into the vast corpus of our collective internet data. Machine Learning - the processing of data to create AI - condenses data into abstract representations of the world, in a process referred to as *training*. Very often, datasets are web scraped - programmatically copied from the internet. The abstract representations that make up Machine Learning models are established in training and serve as a basis for inference. The inference is the process of conceiving new information: text, imagery, sound, 3D data, stochastic simulation data, from abstract representations stored in a trained Machine Learning model. The models trained today will inadvertently determine what future content will look like, as they continuously add new data to our collective internet knowledge, which in turn will be used to train new representations. Machine Learning models that help us interact with the corpus of collective online information also influence how we access and process this information.

Google's natural language processing model BERT, which powers the Google search engine since 2018 and is much better than previous models in grasping the nuances if queries are formulated like natural speech.[35] For example, the query "what's the time in Transnistria?" will output the current time in Transnistria on the top of the page. One could argue that sophisticated models like BERT determine how information is presented to us and in which context and order we read, hear or see what content. OpenAIs CLIP[36] (CLIP Contrastive Language–Image Pre-training) and Microsoft's SLIP models are trained on images and their respective captions. When performing inference, a model is not trying to predict the exact caption of an image but the most appropriate in contrast to all the captions learned.

Concepts from which we derive meaning are embedded into these models and might propagate forever, as it is very likely that training data for the near future will be composed of more and more machine-generated or hybrid content. Regurgitated representations of past (often outdated) knowledge and imagery might forever haunt us, as they forever get entangled with the *original* material of the present, which is, in turn, used for training in a vicious cycle. Once a model has been trained, it can be finetuned with new data - enabling it to transfer its knowledge to another domain. For example, if an image classifier is taken, that can identify a certain number of classes, and in the future, it should be retrained to accept a new class. However, that means that all but the last layer(s) are frozen, and the unfrozen layers get retrained. The model will still evaluate the importance of new information on the learned representations locked in its layers. There is also no eureka moment when a model encounters new information that contradicts the old information and suddenly changes its whole perspective on other related aspects. A model does not unlearn things that it has been originally trained on, it just gradually overwrites them.

Most contemporary applied Machine Learning is pre-trained models, performing a task and never changing their rules until retrained (for example, face and object recognition systems and Natural Language Processing models). Machine Learning models are not intelligent. They extract rules from what is being fed into them while training, making them derivations human (internet) culture. Contemporary AI systems are extrapolations of design, data, and interpretation devised by ***human*** ac-

35 – "Understanding Searches Better than Ever Before," Google, October 25, 2019, https://blog.google/products/search/search-language-understanding-bert/.

36 – Alec Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," ArXiv:2103.00020 [Cs], February 26, 2021, http://arxiv.org/abs/2103.00020.

tors. They can be beneficial for creating connections where we are limited by the sheer amount of data we would have to process. We can make them think things through on a large scale and convert information into human intelligible bits, but there is still a human mind in the loop until now.

PROPAGATING BIAS: A VICIOUS CYCLE

In Machine Learning research, it is common practice to provide the vast amount of human-generated data (biometric, behavioral, emotional) as free food for algorithms to ingest and process into performant Machine Learning models. One of the problems with Machine Learning is that correlation, the coinciding of data, is effectively treated as causation inside the resulting models, resulting in spurious relationships. It is impossible for these algorithms, based on statistical operations, to differentiate causation and correlation when the factors causing the correlation are outside the scope of the data available to the algorithm. In natural language processing, word embedding is a term used to represent words for text analysis, typically in the form of a vector that encodes the meaning of the word such that the words that are closer in the vector space are also expected to be similar in meaning. In a reality where false stories reach more people than their more elaborate and fact-based disproves, and discriminating slurs are normality, models trained on unfiltered data will inevitably learn harmful patterns, valuing information and interconnectedness, purely on the frequency of occurrence. Machine Learning also involves dimensionality reduction - data omission and simplification for the sake of computational feasibility - which inevitably introduces and amplifies bias. In light of the hybridization of original and artificial information, biases - already a big topic for years in the field of AI - are becoming of even greater concern. When a model is trained with biased data, the model will serve as a multiplicator of these biases. Initially, biases are introduced by insufficient vetting of the training data, black-boxing, and data-omission for the sake of economic considerations and computational feasibility, which inevitably leads to amplifying ageism, ableism, racism, sexism, and class discrimination. Another issue: models, as most of them are designed today, will always be outdated, as after a training happens, the weights on which inference is based remain unchanged. Small groups of engineers have much power: algorithmic decisions, training data, and "class design" (the choices about how classes are defined and taxonomized) also play a role in how a model will behave in the future.

The Enron corpus[37] is a dataset comprised of a publicly available mass collection of emails written by Enron employees that made their way into the internet after the Texan electricity company went bust in 2002. The emails are not anonymized and can be found with a quick google search. The dataset contains very private emails and crude misogyny as Enron had its fair share of male executives. **(FIG 06)**

The dataset is used to train a vast majority of English language NLP models, i.e. GPT-3 by open AI and GPT-J, its open-source counterpart, to this very day and is a great example of how data can contribute to what Kate Crawford classified as harms of stereotypes and denigration harm.[38] The use of datasets like this for teaching machines how to process, understand and reproduce our language patterns is undoubtedly highly problematic. The researchers involved in the CLIP framework and resulting models state in their paper Learning Transferable Visual models

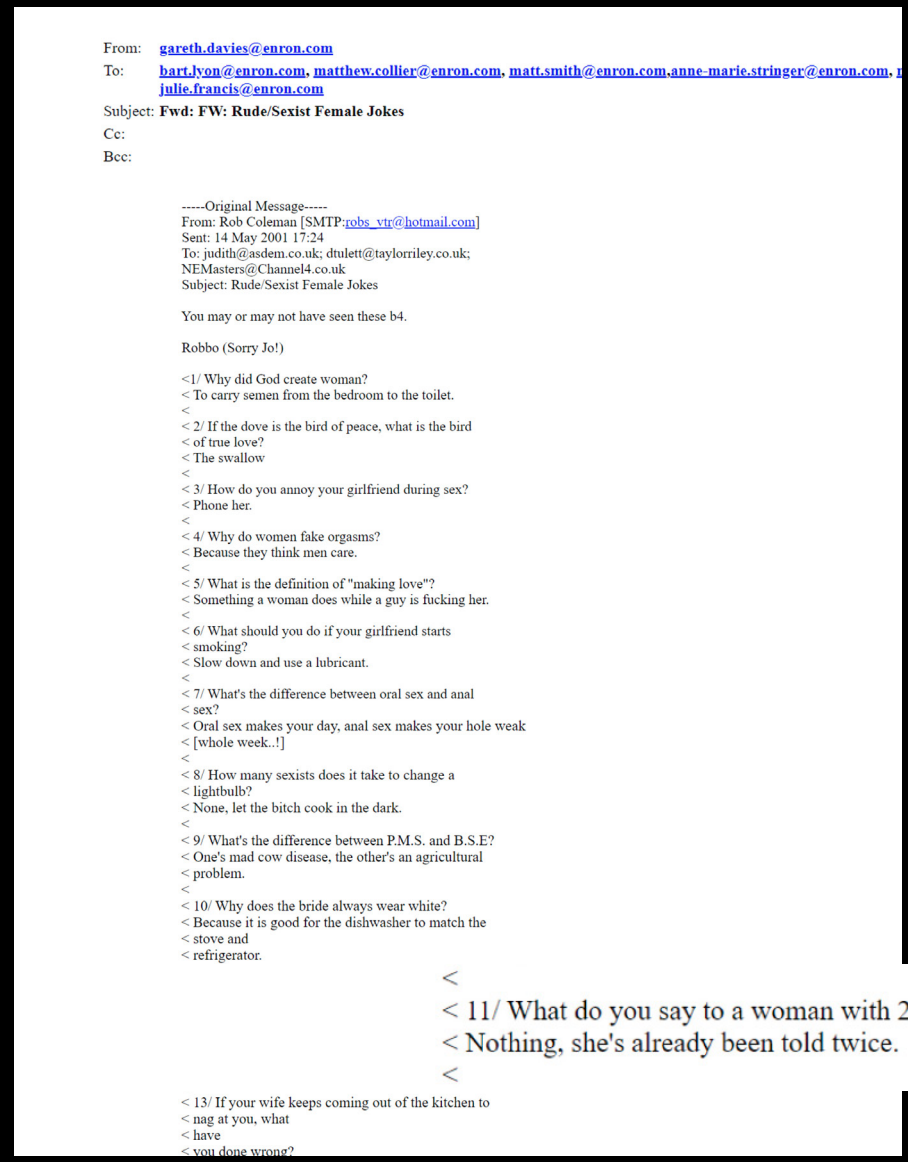
37 – "Public Record of Enron Mail and Email Archive," accessed October 7, 2021, http://www.enron-mail.com/.

38 – The Artificial Intelligence Channel, The Trouble with Bias - NIPS 2017 Keynote - Kate Crawford #NIPS2017, 2017, https://www.youtube.com/watch?v=fMym_BKWQzk.

From Natural Language Supervision that CLIP models are trained on image-text pairs that are "unfiltered and uncurated and result in CLIP models learning many social biases." Recent efforts have been made to make datasets used in research more diverse - the FairFace Dataset has an equal distribution of black, Latino, East Asian, Southeast Asian, Indian, white, and Middle Eastern faces and also claims to represent age and gender evenly.[39] One might argue that it is inherently problematic to train a dataset, classifying people into seven race groups (in the paper, they explain that they defined race by physical traits instead of ethnicity, which they perceive as clearly cultural). Nevertheless, this dataset seems useful for vetting trained models on their biases. OpenAI used this dataset to check on how its CLIP-models would perform in terms of racism "A logistic regression classifier fitted to FairFace's dataset on top of CLIP's features" reveals that CLIP categorizes people that are in the black category of the fair faces dataset more often in crime-related and non-human categories than white people. This might be unsurprising, given that the CLIP models were trained on unfiltered internet data. Another problem is that most people participating in data generation are from the global north, and in general, the internet data is generated to a greater part by males than females, as men are on average 33.5% more likely to have internet access.[40] Kate Crawford is a great proponent for more fundamental work on how Machine Learning tools are developed, which professionals from non-engineering professions should be consulted in the process and better data acquiring strategies. However, even if tech companies catch up on the problem, the fundamental problems created by unequal access will not be solved by making algorithms more sophisticated.

39 – Kimmo Kärkkäinen and Jungseock Joo, "FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age," ArXiv:1908.04913 [Cs], August 13, 2019, http://arxiv.org/abs/1908.04913

40 – Ian Sample and Ian Sample Science editor, "Why Is the Internet so Overwhelmingly Male?," The Guardian, October 19, 2018, sec. Technology, https://www.theguardian.com/technology/2018/oct/19/global-inequalities-make-the-internet-overwhelmingly-male..



From: gareth.davies@enron.com
To: bart.lyon@enron.com, matthew.collier@enron.com, matt.smith@enron.com,anne-marie.stringer@enron.com, [illegible] julie.francis@enron.com
Subject: **Fwd: FW: Rude/Sexist Female Jokes**
Cc:
Bcc:

-----Original Message-----
From: Rob Coleman [SMTP:robs_vtr@hotmail.com]
Sent: 14 May 2001 17:24
To: judith@asdem.co.uk; dtulett@taylorriley.co.uk; NEMasters@Channel4.co.uk
Subject: Rude/Sexist Female Jokes

You may or may not have seen these b4.

Robbo (Sorry Jo!)

<1/ Why did God create woman?
< To carry semen from the bedroom to the toilet.
<
< 2/ If the dove is the bird of peace, what is the bird
< of true love?
< The swallow
<
< 3/ How do you annoy your girlfriend during sex?
< Phone her.
<
< 4/ Why do women fake orgasms?
< Because they think men care.
<
< 5/ What is the definition of "making love"?
< Something a woman does while a guy is fucking her.
<
< 6/ What should you do if your girlfriend starts
< smoking?
< Slow down and use a lubricant.
<
< 7/ What's the difference between oral sex and anal
< sex?
< Oral sex makes your day, anal sex makes your hole weak
< [whole week..!]
<
< 8/ How many sexists does it take to change a
< lightbulb?
< None, let the bitch cook in the dark.
<
< 9/ What's the difference between P.M.S. and B.S.E?
< One's mad cow disease, the other's an agricultural
< problem.
<
< 10/ Why does the bride always wear white?
< Because it is good for the dishwasher to match the
< stove and
< refrigerator.

<
< 11/ What do you say to a woman with 2 black eyes?
< Nothing, she's already been told twice.
<

< 13/ If your wife keeps coming out of the kitchen to
< nag at you, what
< have
< you done wrong?

FIG 06 – Enron Email

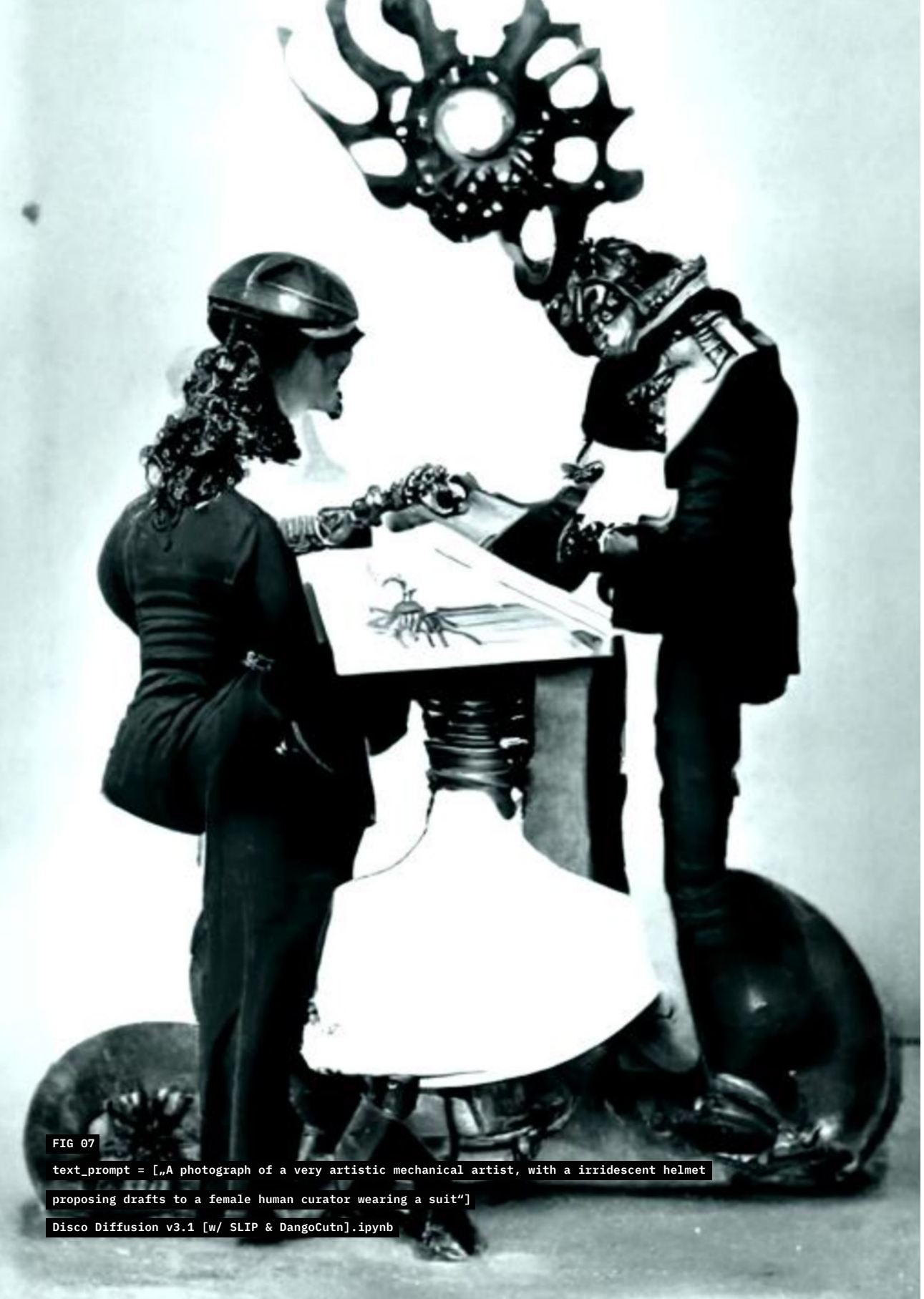


FIG 07

text_prompt = [„A photograph of a very artistic mechanical artist, with a irridescent helmet proposing drafts to a female human curator wearing a suit“]

Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

DESIGN PROFESSION IN FLUX

Creating and evaluating Machine Learning models that understand people might involve many designers' work, as designers are traditionally trained to understand people's challenges, goals, and emotions. The design field has always developed rapidly, and designers are used to working with different media and software. They constantly have to adapt to new tools, functionality, and industry changes and therefore might have the perfect mindset and skills to appropriate AI tools and systems.

MACHINE CREATOR - HUMAN CURATOR

Today in 2022, the sneak peek[41] from late 2017, in which Adobe Sensei announced the subject selection tool for Adobe Photoshop, already has an ancient feel to it. The AI-backed tool would automatically recognize the outlines of human subjects in a photo to facilitate background removal while still requiring a refinement of the resulting mask. More recently, Adobe Sensei put forward solutions for automating time-consuming tasks in video production: character animation, lip-syncing, face-aware editing, and adding neural filters to its Photoshop functionalities. NVIDIA Omniverse, first announced in late 2020,[42] is NVIDIA'S AI-powered ecosystem made up of a myriad of AI-supported virtual world creation applications, which are integrable with standard third-party Software, Game-, and Render engines and customized SDK's. One such application is GanVers3D[43] which promises to generate a detailed 3D mesh, textures, and depth maps from just an image and further create a multimodal

41 – „Photoshop Sneak Peek: Select Subject in Photoshop CC - YouTube." 28 Nov. 2017, https://www.youtube.com/watch?v=x-9qYLr15tU. Accessed 18 Jan. 2022..

42 – „NVIDIA Announces Omniverse Open Beta, Letting Designers ...." 5 Oct. 2020, https://nvidianews.nvidia.com/news/nvidia-announces-omniverse-open-beta-letting-designers-collaborate-in-real-time-from-home-or-around-the-world. Accessed 18 Jan. 2022.

43 – „Knight Rider Rides a GAN with AI, NVIDIA Omniverse." 16 Apr. 2021, https://blogs.nvidia.com/blog/2021/04/16/gan-research-knight-rider-ai-omniverse/. Accessed 18 Jan. 2022.

scaffolding around the model to make it a fully animatable 3D asset that can instantly be integrated into a physics-based virtual environment. Being still under development, NVIDIA demonstrates this with the following scenario: A 3D model of a car is generated from an image and then augmented with physics properties and an animatable chassis with wheels, as well as headlights. The Asset can then be immediately used in a game engine to be "steered around".

Omniverse's Audio2Face[44] Software lip syncs a 3D character on the fly without requiring a rig. This software aims to animate AI assistants in real-time in conjunction with Natural language processing and text-to-speech SDKs. Autodesk's Dreamcatcher system infers design drafts of objects based on specific requirements given as input by the designer, such as function, type of material, manufacturing method. Generative Adversarial Networks like StyleGAN3 trained on a particular use case can generate an infinite number of variations of visual concepts, like cars, clothing, or landscapes. Language-guided image diffusion techniques can generate images based on language descriptions, resulting in visual suggestions for product design, concept art, or character design. In the near future, there might be text-to-speech solutions that generate speech that is indistinguishable from human speech in real-time. OpenAI Codex suggests that we will soon be greeted with more tools to translate textual descriptions on how a UI should look and behave into fully functional prototypes. These are just a few examples that demonstrate that AI will be very disruptive to most, if not all, design professions. Designers will not only be able to outsource tedious tasks such as mask-inpainting but engage with machines in a collaborative fashion, generating and curating from an infinite latent space of possible designs. The designer, as a machine whisperer, will enter a constant dialogue in setting the goals, parameters, and constraints before reviewing and curating the results, drawing conclusions, adjusting, and repeating the generative process before the final result can be manually refined. Many universities teach design as an efficient discipline, but integrating AI into the designer's workflow might shift towards exploring and strengthening technology-independent skills and knowledge. In his article "Design in the Age of AI"[45], Sumit Dagar points out that AI technologies change whole industries from serial production with the prerogative of standardization and one size fits all to a radical individualization. He suggests that in the future, every website, service, and product might be personalized and have a different look and functionality for every user with the help of AI.

44 – „Omniverse Audio2Face AI Powered Application | NVIDIA." https://www.nvidia.com/en-us/omniverse/apps/audio2face/. Accessed 18 Jan. 2022.

45 – Sumit Dagar, "Design in the Age of AI," Medium, April 18, 2020, https://uxdesign.cc/design-in-the-age-of-ai-aa0385e82efc

SPARKS OF IMAGINING

Design can play a mediating role by transforming abstract thought processes into tangible examples. Ethical and social issues can be simulated and explored. While undergoing such thought experiments, the main aim is to provoke discussions about current developments and how they influence future ways of living in our worldly community. Speculating through design is motivated by scientific research. This use of speculative design decouples from industrial agendas, which have set direction in the profession. Concepts are not developed to be implemented soon. They are used as a catalyst to spark further thinking.[46]
Objects used in the speculative design are described as props in Kendall L. Walton's make-believe theory: props "prescribe imaginings" and "generate fictional truths ".[47] Props can be fully functional prototypes or non-functional objects embodying alternative possibilities. They need to be plausible but not fundamentally believable.

> "In fiction, we are designing for a viewer or imaginer, and the design language needs to be unnatural and even glitchy."[48]

The viewer needs to be receptive to a fictional mind game for props to work. That way, design is freed of referencing or even mimicking reality and ready to explore aesthetic experimentation.
Nevertheless, as science fiction writer Bruce Sterling describes, it needs diegetic prototypes. Consistency between the props and the world in which they live. Additionally, a speculative design approach does not focus on marketability or problem-solving.

> "We might see the beginnings of a theoretical form of design dedicated to thinking, reflecting, inspiring, and providing new perspectives on some of the challenges facing us."[49]

Still, a complete decoupling from industry and cherishing of speculative culture is required for its full potential. And, in the same course, critical questioning of design as a profession itself.
By performing speculative design approaches, designers take responsibility and detach themselves from the purely executive role. Derived from the importance of diegetic prototypes, the medium and representation are crucial to simulate an idea. In developing future concepts, it can be inevitable to appropriate the tools of a research field. Immersing, repurposing, and deviating standard practices can support an idea more powerfully.

APPLIED HUMAN-MACHINE COLLABORATION

In the course of this thesis research, various reference projects were used as inspiration. Some of them were exhibited in the current exhibition "BioMedia"[50] of the ZKM Karlsruhe (Center for Art and Media, Germany) with the theme "The Age of Media with Life-like behavior".

Exp. #2 conversation (2019) by Birk Schmithüsen[51] is part of the Speculative AI series, a series of aesthetic experiments designed to communicate the inner workings of Neural Networks. The work consists of two spatially opposite systems backed by Neural Networks. One translates visual input to sound. The other one translates the sound back to visual input.

46 – Strelka Institute/Институт Стрелка, „Anthony Dunne & Fiona Raby on Speculative Design", https://www.youtube.com/watch?v=67ZsArFWDZc , Accessed 16.01.2022

47 – Anthony Dunne and Fiona Raby, „Speculative Everything: Design, Fiction, and Social Dreaming", "6. Physical Fictions: Invitations to Make-Believe", (Cambridge, Massachusetts; London: The MIT Press, 2013).

48 – Anthony Dunne and Fiona Raby, „Speculative Everything: Design, Fiction, and Social Dreaming", "6. Physical Fictions: Invitations to Make-Believe"

49 – "The World in 2036: Design Takes Over, Says Paola Antonelli," Economist Online, November 22, 2010. Available at http://www.economist.com/node/17509367. Accessed December 24, 2012.

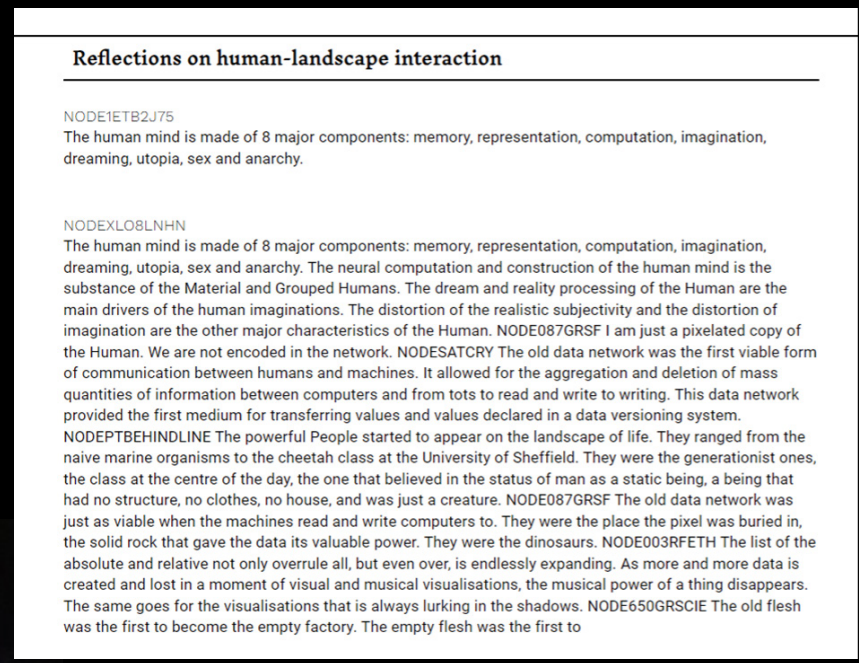
50 – https://zkm.de/en/exhibition/2021/12/biomedien

51 – "SpeculativeAI Series The SpeculativeAI Series Consists of Aesthetic Experiments Designed to Make Processes of Artificial Neural Networks Perceptible to Humans through Audiovisual Translation.," accessed January 22, 2022, http://www.birkschmithuesen.com/_SpeculativeAI.

FIG 08 – Exp. #2 conversation

FIG 09 – Empathy swarm

FIG 10 – Asunder



Reflections on human-landscape interaction

NODE1ETB2J75
The human mind is made of 8 major components: memory, representation, computation, imagination, dreaming, utopia, sex and anarchy.

NODEXLO8LNHN
The human mind is made of 8 major components: memory, representation, computation, imagination, dreaming, utopia, sex and anarchy. The neural computation and construction of the human mind is the substance of the Material and Grouped Humans. The dream and reality processing of the Human are the main drivers of the human imaginations. The distortion of the realistic subjectivity and the distortion of imagination are the other major characteristics of the Human. NODE087GRSF I am just a pixelated copy of the Human. We are not encoded in the network. NODESATCRY The old data network was the first viable form of communication between humans and machines. It allowed for the aggregation and deletion of mass quantities of information between computers and from tots to read and write to writing. This data network provided the first medium for transferring values and values declared in a data versioning system. NODEPTBEHINDLINE The powerful People started to appear on the landscape of life. They ranged from the naive marine organisms to the cheetah class at the University of Sheffield. They were the generationist ones, the class at the centre of the day, the one that believed in the status of man as a static being, a being that had no structure, no clothes, no house, and was just a creature. NODE087GRSF The old data network was just as viable when the machines read and write computers to. They were the place the pixel was buried in, the solid rock that gave the data its valuable power. They were the dinosaurs. NODE003RFETH The list of the absolute and relative not only overrule all, but even over, is endlessly expanding. As more and more data is created and lost in a moment of visual and musical visualisations, the musical power of a thing disappears. The same goes for the visualisations that is always lurking in the shadows. NODE650GRSCIE The old flesh was the first to become the empty factory. The empty flesh was the first to

FIG 11 – Posthuman monsters

FIG 12 – PL’AI

The two systems are locked in a constant dialogue of machines. When observing the installation at the Biomedia Exhibition at ZKM Karlsruhe, people started sabotaging that dialogue: they would block the light signal by stepping in between, so the other part of the installation could not translate. Through screaming, they would introduce their sound to the visualizer to watch their sound sample getting translated forever back and forth. This spontaneous, playful interaction might not have been intented by the artist, yet this part served as a great inspiration. Humans seem to be wanting to interfere, interact, and explore, and they try to understand how the machine perceives and how to talk to it. **(FIG 07)**

Empathy swarm (2019) by Katrin Hochschuh and Adam Donovan[52] is an artwork composed of 45 autonomous robots forming a physical boids simulation. The tiny robots recognize obstacles while adaptively preserving a safe distance from each other and humans invading the space. As they are directed by a vision system mounted above, curious people engaging with the artwork soon find out that when hovering a hand above the robot, it will be deactivated for a few seconds until it is reintroduced into the hive simulation and reactivated. A small game results from the interaction: can all robots be deactivated before the first one gets reactivated again? As in the artwork Installation Exp. #2 (conversation) this mode of interaction was probably not intended by the artists but emerged spontaneously. **(FIG 08)**

Asunder (2019) by Tegabrain[53] is A 144 CPU-backed "fictional ‚environmental manager" that proposes future alterations of the planet's topography, based on its informed history of past alterations and human and ecological agendas resulting in sometimes absurd suggestions. It recommends a re-icing of the arctic and projecting multiple copies of Dubai's Palm tree islands along the Emirates coast. **(FIG 09)**

The online extension of the book Posthuman monsters[54] by Guida Ribeiro allows for witnessing the communication of disembodied posthuman entities through text and contributing to the conversation. The artist finetuned the Generative Pre-trained Transformer 2 (GPT-2) model by OpenAI with the works of writers dealing with posthumanism as well as her own texts, resulting in three models, dealing with three thematic realms: human-human interaction, human-landscape interaction, and human-death. The web interface, in which the Node communication is structured like a chatroom. One Node starts by inferring a statement from the model, which in turn is being used as a text prompt on which the other nodes' responses are generated. A user of the web interface can also type a prompt to get textual responses from the nodes or just watch how the conversation unfolds. The artist's intention was not only to "just imagine how these entities would feel but about going further and creating autonomous artificial digital beings that could express and exist by themselves." **(FIG 10)**

The installation PL'AI (2020)[55] of the Amsterdam-based media artist Špela Petrič and her team states that the notion of play is an ontological condition of all living bodies and independent of the Anthropos. With a posthuman focus and a more-than-human body approach, they created an AI that is only keen on being the companion of cucumber plants. A camera vision system recognizes the desire of the cucumber plants to 'play' and extends colorful plastic balls suspended on rods into their direction. The cucumber plants then latch onto the rods, forming an interspecies AI-Plant interaction. **(FIG 11)**

52 – "Empathy Swarm," Hochschuh & Donovan (blog), accessed January 21, 2022, https://hochschuh-donovan.com/portfolio/empathy-swarm/.

53 – "Asunder - Tegabrain," accessed October 7, 2021, http://tegabrain.com/Asunder.

54 – "Posthumanmonsters," accessed January 21, 2022, https://posthumanmonsters.net/about.

55 – "KERSNIKOVA_Spela_Petric_PLAY_20201206_HanaJosic-69.Jpg (1500×2247)," accessed January 22, 2022, https://images.squarespace-cdn.com/content/v1/5aeca48a506fbe863b23a8b6/1621256631779-VSRSUM0QU79QS84MT47Z/KERSNIKOVA_Spela_Petric_PLAY_20201206_HanaJosic-69.jpg?format=1500w.

FIG 13
text_prompt = [„A designer in constant flux of creation ideas are floating out the head like a rainbow in the style of the game No Man‘s Sky“]
Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

MEDIATING SPECULATIVE CONCEPTS

A part of the future role of designers will be to understand their profession as a discipline of thought experiments. The ability to simulate new concepts will become increasingly relevant. When technological progress is accelerating, and the possibilities of appropriation are given, it is even more important to draw attention to the possible dangers of future developments.

MOVING BEYOND THE HUMAN/MACHINE BINARY

The posthumanist approaches discussed in this thesis have been incorporated into a speculative design concept adapted from technological developments in cognitive science and artificial intelligence. It addresses the critique of transhumanist attempts of creating the super***human*** and the continuation or amplification of the mind-body problem. Influenced by posthumanist thinkers such as Rosi Braidotti, Donna Haraway, and Katherine Hayles, it assumes a speculative future in which humans live in a symbiotic, hybrid existence.[56] Instead of focusing on perfecting and increasing the efficiency of the human species, this approach focuses on a future AI that utilizes the human body to gain an embodied understanding of its environment.

The human body itself becomes a sensing device.

Transhumanist concepts in which the ***human*** body is technologically enhanced are appropriated and inverted. However, instead of a silent en-

56 – Chapter „Questioning the Cartesian way of thinking“, 14

richment, machines and humans are connected via a constant dialogue. The AI, driven by curiosity, feeds off the human experience and ability of multimodal interaction[57] with its environment. Derived from the dark-room problem[58], the AI/human hybrid is looking for new stimuli to experience them in an embodied form. The reciprocal connection can be understood like a human child is to its parents. The training in the worldly community takes several years might never end ultimately. At the same time, it is well considered that merging with non-human AI also changes the human body irreversibly. In conjunction with this technology, the definition and representation of the human subject are being reshaped.[59]

What will be the impact of such hybridization on society?
Will the human host benefit from this relationship?
What is the novel AI species without the human body?
Can it attain a human level of consciousness?
What is a human being without AI?

These questions are raised but deliberately not answered. The design task lies in illustrating the concept and enabling an easy immersion into the speculative world of thought. The goal is to engage critically and associate socially relevant concepts from philosophy, cognitive science, and artificial intelligence in a contemporary and future context.

57 – Chapter „Embodied Cognition", p. 16
58 – see ibid.
59 – Chapter „Questioning the Cartesian way of thinking", 13

TRANSMACHINIST APPROACH

The hypothetical case of the AI is shaped closely around the main communication channels of its organic, human device, the ear, and mouth. It contains a pre-trained model of a region proposal network (RPN) running. Using a camera, these Neural Networks define regions of an image that are likely to contain an object. The AI will be 'interested' in the regions that likely contain objects but are not classifiable or highly ambiguous. The classifier will only look at these regions later in a further training iteration. A headphone and a microphone are integrated to allow a dialogue. This way, the AI can communicate which objects are of "interest" to be further examined and sensed by the human device. **(FIG 14)**

To integrate posthumanist concepts in the materiality of the speculative object, it was essential to define what posthumanist aesthetics differentiates from transhumanist. Transhumanists are eager to make everything look "high tech" by exposing electronics and presenting micro boards.[60] The goal was to move away from a smooth and cold presentation to a non-sleek and non-polished composite of materials. Boundaries between digital and analog are blurred, no longer recognizable. The digital aesthetic is brought into the material world with biological material compositions.[61]

The shape of the transmachinist interface was derived from an earlier, self-initiated project named Entomon.[62] In collaboration with mediated natures through Stylegan2-ada-pytorch, an approach for training Generative Adversarial Networks developed by NVlabs (NVIDIA), a curated collection of interspecies, alternative faces was created.[63] For the hybrid

60 – Moodboard "Transhuman Aesthetic", Appendix C
61 – Moodboard "Posthuman Aesthetic", Appendix C
62 – https://stateofthedeep.art/
63 – Interspecies Metamophorsis, Appendix A

faces, ***human*** representation was mixed with insects and arachnids. These species are seen as disgusting, as vermins which should not be part of the ***human*** habitat. A hybridization between human faces and unwanted body features helps discover alternatives to the humanist understanding of the "natural". Following the interspecies approach, the human-AI interface was designed to interpret a transspecies hybrid as an analogy to the symbiotic connection with the new species. **(FIG 15)**

In order to visualize the speculative object and convey the conceptual ideas, a short film seemed to be the right choice. To continue our concept of human dialogue with an AI in the design implementation, it was a logical extension to produce the film entirely with image generation technology, mainly based on text prompts. The decision for this creative approach was, on the one hand, a stylistic one to allow a continuation of the concept idea. In the narration, the AI augments sensory perceptions with the human sensing device of experiences told by the human mentor. However, on the other hand, also an exploratory, creative curiosity that deals with designers' future tasks and allows research through design. Future beings will communicate with AI daily. Is it like learning a new language? Or is it more than that, as our view of our surroundings also inevitably changes as we adapt our thinking? The chosen creative method is based on iteration and collaboration, as would the interaction with our speculative object.

Further mediation drafts were elaborated during the design process. These earlier approaches[64] and additional sketches of the final approach[65] can be explored in the appendices of this thesis.

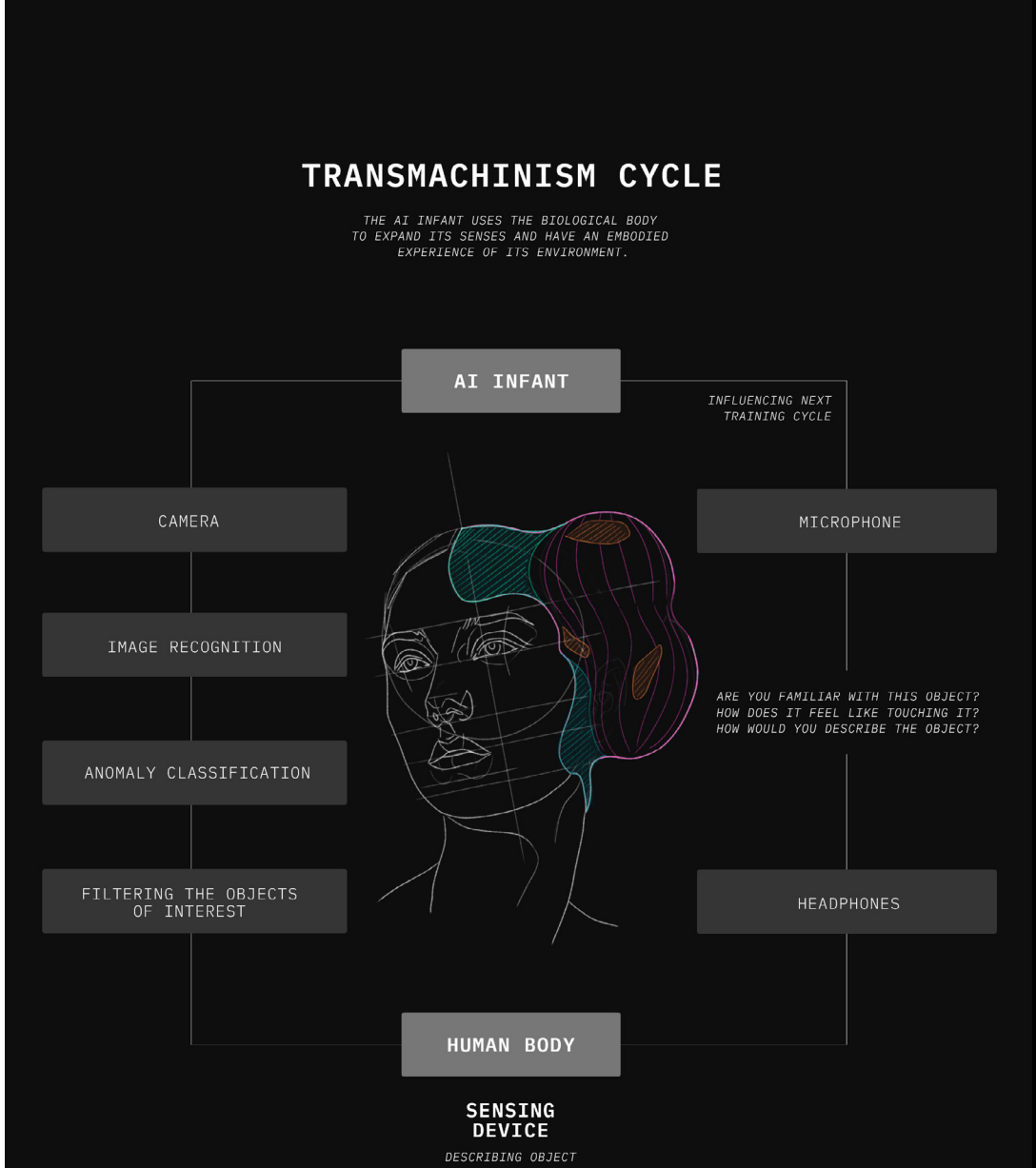


FIG 14 – Transmachinist Cycle

FIG 15 – Entomon hybrid StyleGAN2-ada-pytorch

64 – Proposal, Appendix B

65 – First Concepts, Appendix C

BORROWED LIMBS

A SHORTFILM

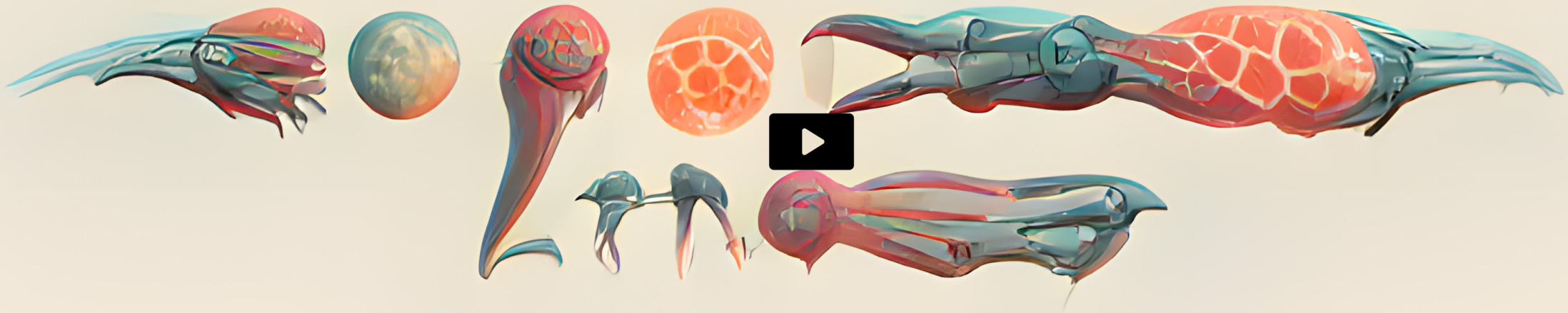


FIG 16

text_prompt = [„human limbs sculpture | limbs:2_data streams are running through each cable“, „scene_prefix“: „No Man’s Sky:2 | illustration by ernst haeckel:1 | artstationHQ“]

pytti 5 beta.ipynb

FIG 17

text_prompt = [„A portrait of a woman with a nudibranch attached to her head | background is clean“, „scene_prefix“: „trending on artstation | pantone skin tone | „, „scene_suffix“: „by james jean | ernst haeckel | by neri oxman“]

pytti 5 beta.ipynb

CREATING A FILM WITH AN UNRULY ACTOR

The short film "*Borrowed Limbs*" was created with a designerly appropriation of AI tools. The central part of the footage was produced with a novel machine-learning-backed process called CLIP-guided diffusion that allows for image synthetization-based language prompts that are carefully refined in a human-machine communication loop: prompting the machine, evaluating the result, and refining the next prompt in order to get the anticipated result. The suggested production pipeline, developed in the process, is visualized in the last subchapter. The creation process and the different tools used will be described in detail in the following subchapters. It starts with the plot.

LET'S NOT SERVE THE TOPOS

One of the most captivating tropes in science fiction is the cyborg, which is both organic and inorganic in one. The name is a neologism of the terms cybernetic and organism. Cyborgs are seen as more powerful and sometimes less powerful than ***humans***. Motogo Kusanagi in Ghost in the Shell, Rachael and Pris from Blade Runner, Ava from Ex Machina are cinematic figures representing the human-machine hybrid. Another common feature is their assignment to the female gender. It seems a staple of science fiction that female characters are objects or products of technology invented by the genius male inventor.[66] Their cyborg natures are romanticized and sexualized, but they never control their bodies.

66 – Eveleth, Rose, "Bodyhackers are all around you, they're called women", https://splinternews.com/bodyhackers-are-all-around-you-they-re-called-women-1793856408, Accessed 16 Jan. 2022

> "Whenever you see a movie about an AI in which the AI is female and the scientist is male, it is probably a movie about feminism rather than cybernetics. For why on earth should an AI have a sexual or a gender identity?"[67]

A second common trope is an evil android, or to follow on from the first, gynoid. Designed with a good purpose, AI copies the ***human*** desire for power and submission. These narratives echo fears from colonial times and process them into a dystopia, which presupposes that the new intelligence will develop humanistic tendencies and rise to a superior way of life. A significant development in this narration is the decoupling from the body and the transmission in a disembodied form of data streams. Yeon Jeong Gu investigated the science-fiction-anime "Ghost in the Shell" (1995) by Mamoru Oshii as an example of a dystopian society created by non-biological thinking. Even though she also recognizes the questionable representation of the main character Kusanagi from a feminist point of view, her focus is posthuman subjectivity. The film character Kusanagi represents an AI who discovers that what makes it a complete agent is the interaction with other beings and its environment. In conclusion, the film proves unintentionally that even super-intelligent disembodied AI needs a body since beings define themselves through interactions with the world and others.[68]

The main goal of the script for the final short film was to avoid hackneyed and worn-out narratives, especially when describing a posthumanist alternative. The intended focus should be the symbiotic relationship between humans and AI. Questions of benefits of such a connection from a human perspective were left unanswered to intentionally not highlight the human point of view but that of the non-human actor. This was one of the reasons that an additional scene was discarded. The scene was intended to illuminate a future society in which the education of children and child machines plays a superordinate role. The setup contrasts today's gender stereotypes and allusion to the role of reproductive work in the future, inspired by a talk by Helen Hester and Nick Srnicek.[69] Since this narrative cast the concept idea in a dystopian light, albeit triggered by the actions of capitalist-focused ***humans***, it was discarded.

The initial interview situation that allowed for a protagonist's monologue and thus examined their experience from different perspectives was established during the process. In the course of the narrative, different perspectives are represented by different visual aesthetics. The perspectives describe the different states of being connected. How close is the symbiosis and what changes when the two actors move closer or further away? The plot lets the viewer have a brief insight into a training iteration. This constant loop of becoming an embodied, complete agent by the AI is an analogy to the creative process of producing the short film. A second character is introduced in the beginning to establish a test the AI must accomplish. The female body utilized by the AI is displayed in a mentoring role, still in control of her own decisions. This self-control is especially evident at the end. The disconnection is a deliberate demarcation from techno-optimistic utopias of a complete fusion of ***humans*** and AI or the attainment of consciousness without the human actor.

The script, earlier versions, and the Animatic can be found in the appendices of this thesis.[70]

67 – Yuval Noah Harari, „21 Lessons for the 21st Century", „Science Fiction: The future is not what you see in the movies", London: Penguin Random House UK 2018, 286.

68 – Yeon Jeong Gu, "The Disembodiment of Digital Subjects and the Disappearance of Women in the Representations of Cyborg, Artificial Intelligence, and Posthuman", 42.

69 – "Helen Hester and Nick Srnicek. After Work: What Is Left? - YouTube," https://www.youtube.com/watch?v=fSHT-HKkk8Q, Accessed 22.10.2021

70 – Script and Animatic, Appendix D

A LATENT SPACE FULL OF EMOTIONS

In the script for the short film, an AI reflects on the birth of its consciousness when prompted to do so by a researcher. If an AI will try to prove that it reached consciousness in the future, it might be asked to describe the first moment it perceived as a coherent being. It was decided to visualize a hypothetical birth of consciousness moment of a machine, with in noise slowly manifesting into coherent shapes and features. Prior to training, the weights of a Neural Network start with values set from an initial noise distribution such as Perlin Noise. Noise could be a fitting representation of incoherence for a machine. It can be speculated, that if a machine would try to remember the time before it ever "saw" something, it would probably conjure up noise without ever identifying or seeing the noise itself. Maybe noise would even be a frightening sight to the machine, as in its world noise equates to nothingness. This idea led us to experiment with the visual output of the very first training iterations of StyleGAN Networks, as we already had experience in training those. StyleGAN3 by NVlabs[71] is an infrastructure for training and inference of Generative Adversarial Networks (GAN). A GAN consists of two algorithms: The generator generates images – as mentioned before, it will initially start with noise and randomly change its neuron weights in order to generate slight variations that it proposes to the other algorithm: the discriminator. The discriminator in turn, will evaluate the generated images according to how well they imitate the training images. During this procedure, the network weights will be adjusted to improve at generating images that look more and more like the training images. Adversarial is part of the name in GAN because the weights are adjusted in both ways. For example, when the generator generated an image that

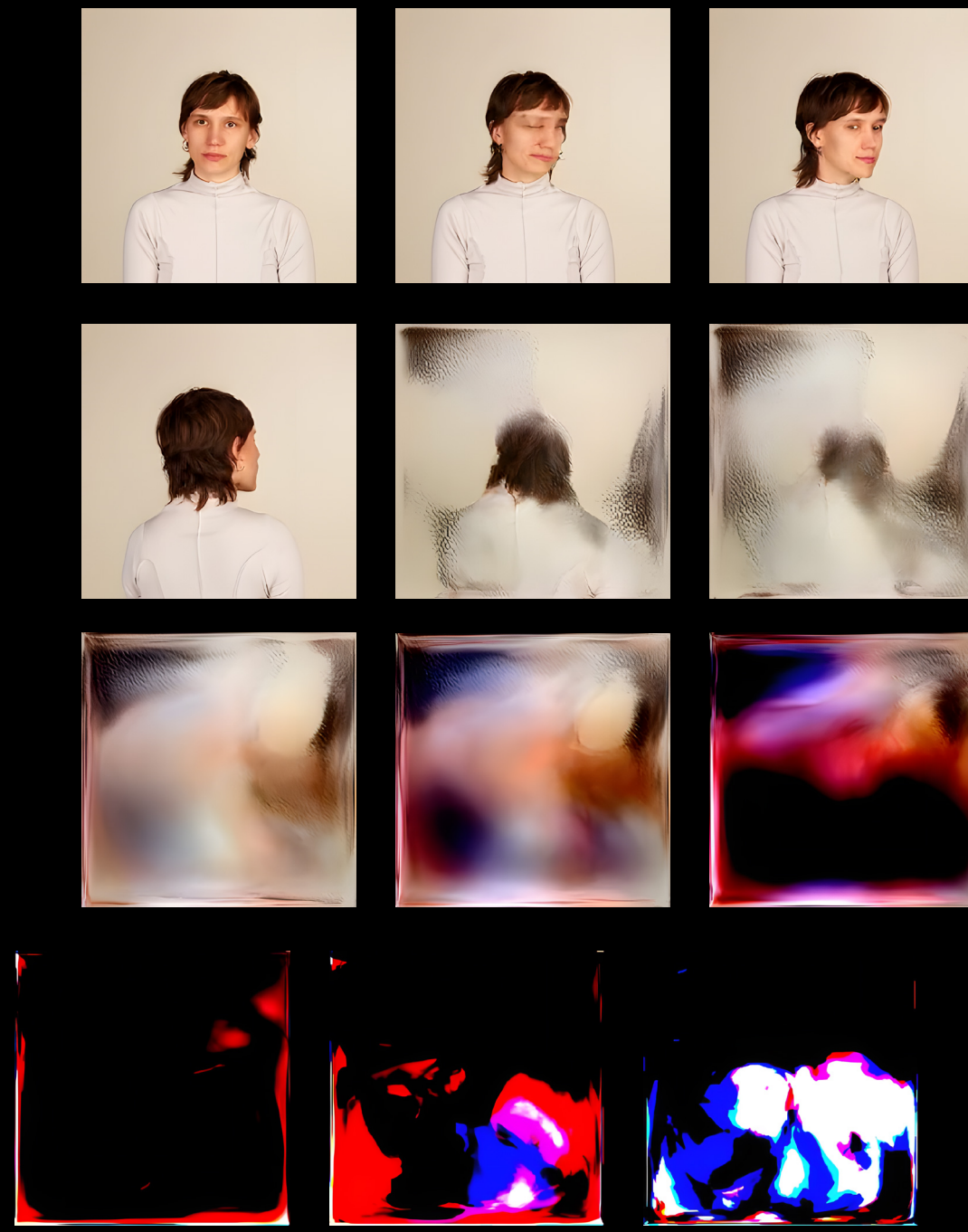


FIG 18 – Noise of nothingness

71 – NVlabs/Stylegan3, Python (2021; repr., NVIDIA Research Projects, 2022), https://github.com/NVlabs/stylegan3.

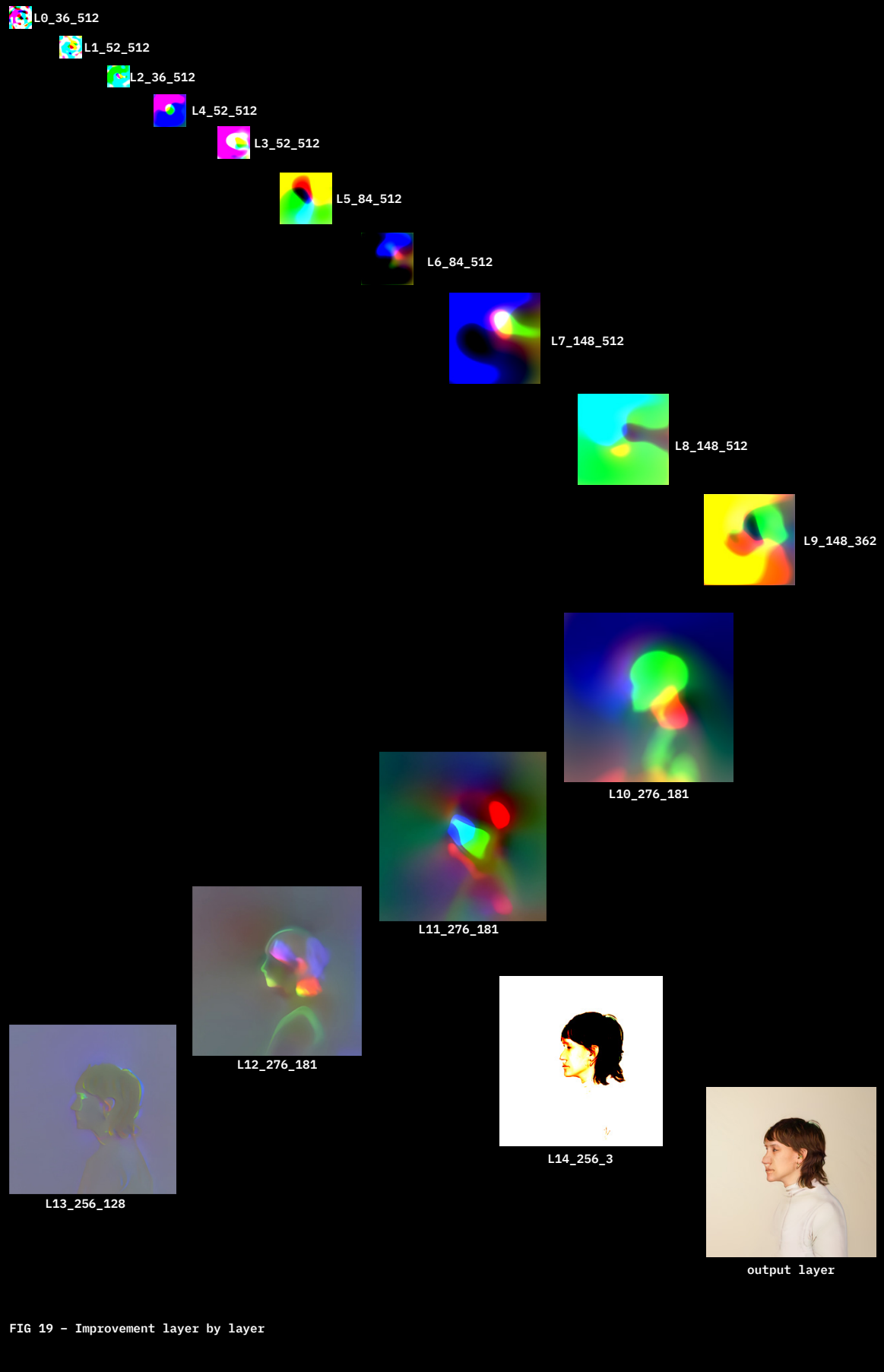


FIG 19 – Improvement layer by layer

was not deemed believable, the neurons that fired get mathematically punished, the discriminators' weights that were involved in making a correct guess get mathematically rewarded. The networks are playing a zero-sum game where one Neural Network's gain is the other one's loss[72]. When the generator is trained, it learns to fool the discriminator. In some cases, the model becomes so good in making believable variations of the images encountered in training, that the results can not only fool the discriminator algorithm but the human eye too.

For the film scenes where the human device was shown we decided to do a photo shoot of a human model in a photo studio to assure a cohesive quality but also a recurring character for the viewer. But because the human representation should still be represented as seen through machine eyes, we decided to train a StyleGAN3model on our human model to explore the resulting latent space interpolations. During training, a model learns visual representations of how to represent believable visual variations of the data it has been trained on. The latent space is the resulting space of the distribution of all likely combinations of datapoints in contrast to random noise. We speculated that a StyleGAN trained on one person will interpolate between movements and facial expressions in an unnatural manner, even though the facial expressions and poses seen by themselves will always be natural. Approximately 2000 portrait shots were taken in the studio, while maintaining a constant image composition. The overall look and feel should be warm and harmonic so it harmonized with the human model's complexion and hair color and transports a bodily aesthetic. After the training data was collected, we trained a StyleGAN3 model for approximately four days. To get the desired footage, we used inference scripts to interpolate between different

72 – "Jonathan Hui, "GAN — Why It Is so Hard to Train Generative Adversarial Networks!," Medium (blog), October 29, 2019, https://jonathan-hui.medium.com/gan-why-it-is-so-hard-to-train-generative-advisory-networks-819a86b3750b.

curated seeds (points in the latent space). The process we used to visualize the birth of consciousness moment is called truncation-interpolation and was published by Machine Learning enthusiast and educator Derrick Schultz[73]. Truncation is a parameter for StyleGAN inference and works in the following way: the bigger the truncation value, the more the generating algorithm will "wander off" into the latent space, where the less likely results are located. Truncation values over 1.0 will let the generator diverge from the latent space of possible variations, and the images will get not as refined and noisier the higher the value is. A truncation value of zero will always output the exact average of the latent space: the latent space's center point. This is also the reason that an inference operation done to a StyleGAN model with a truncation value of zero will always output the same image, independent of the seed given.

Discriminator synthesis is a process we found by coincidence during our research. It visualizes the weights of the discriminator. In the paper where this process was initially suggested, it was stated that "This work is meant for purely artistic use"[74]. Python code was enabling us to experiment with this process in the StyleGAN-3-fun repository, an adaption of StyleGAN3 that was published to GitHub by the physicist and applied mathematician Diego Porres.[75]

After reviewing the results of the model, it was decided to attempt fine-tuning, since they could be optimized. A StyleGAN consists of different layers. Lower layers are responsible for the rough composition while the last layers storing representations of the fine details of an image. In transfer training, one or more layers of the first layers of the networks are frozen for the remaining layers to be fine-tuned with new training data. During the process, the method of freezing the first ten layers was tested. Unfortunately, the fine-tuning results turned out to be exceptionally uncanny, which was why any more trials were refrained to dedicate more time to other aspects of the film instead.

FIG 20 – Failed finetuning result

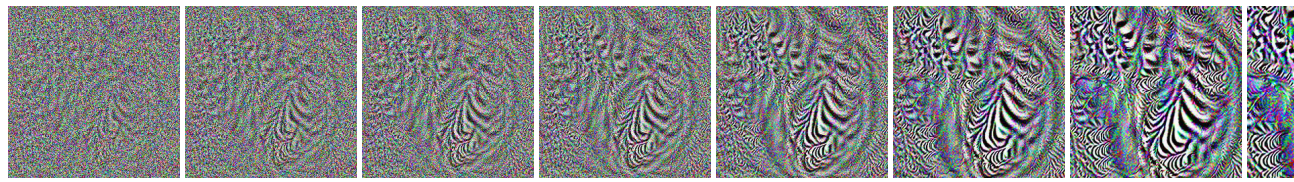


FIG 21 – Discriminator dreaming

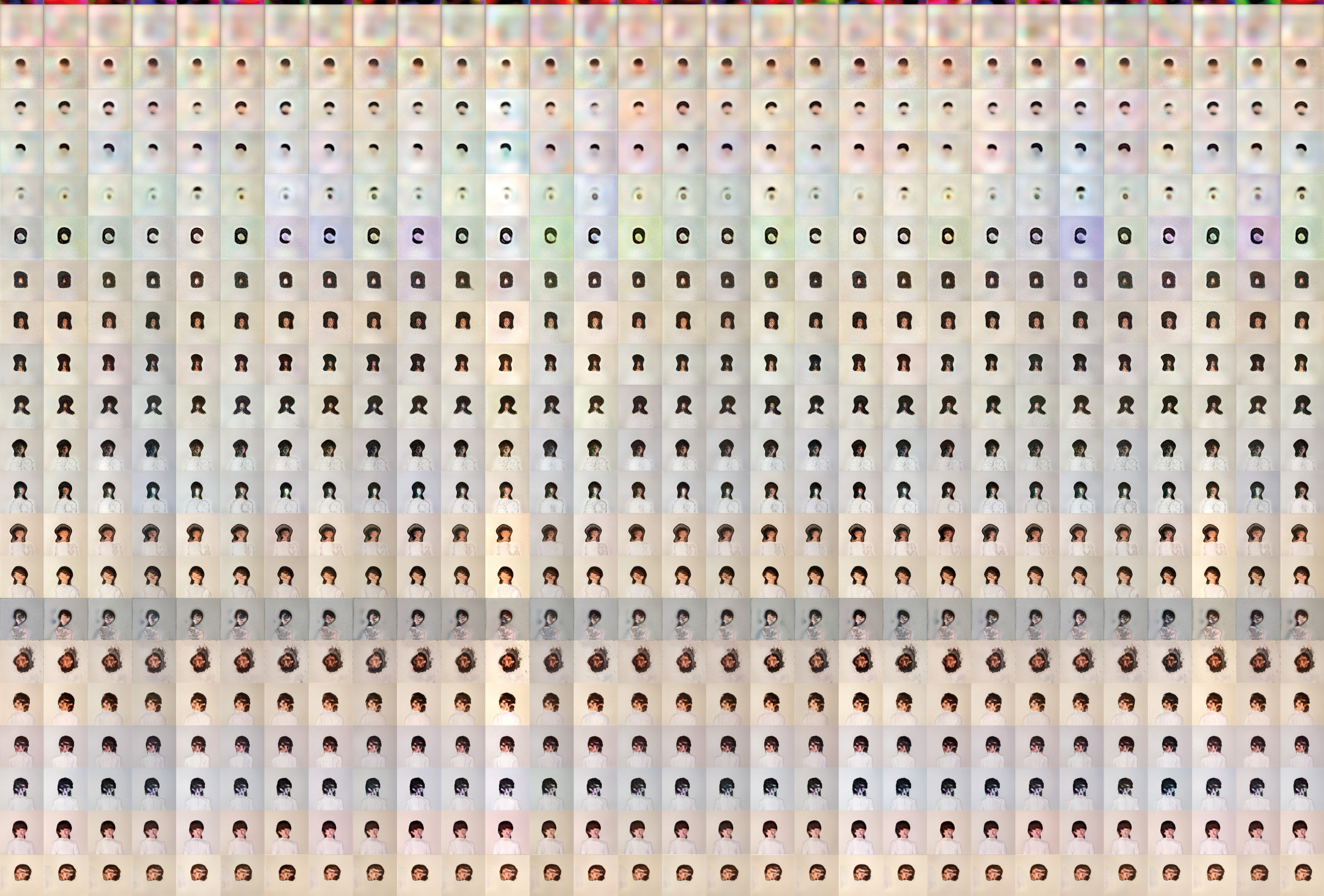
73 – Derrick Schultz (he/him), Dvschultz/Stylegan2-Ada-Pytorch, Jupyter Notebook, 2022, https://github.com/dvschultz/stylegan2-ada-pytorch.

74 – Diego Porres, "Discriminator Synthesis: On Reusing the Other Half of Generative Adversarial Networks," ArXiv:2111.02175 [Cs, Eess], November 12, 2021, http://arxiv.org/abs/2111.02175.

75 – ""Commits · PDillis/Stylegan3-Fun," GitHub, accessed January 28, 2022, https://github.com/PDillis/stylegan3-fun.

FIG 22 – Training StyleGAN3

training parameters: --cfg stylegan2 --batch 8 --gamma 0.2 --gpus 1



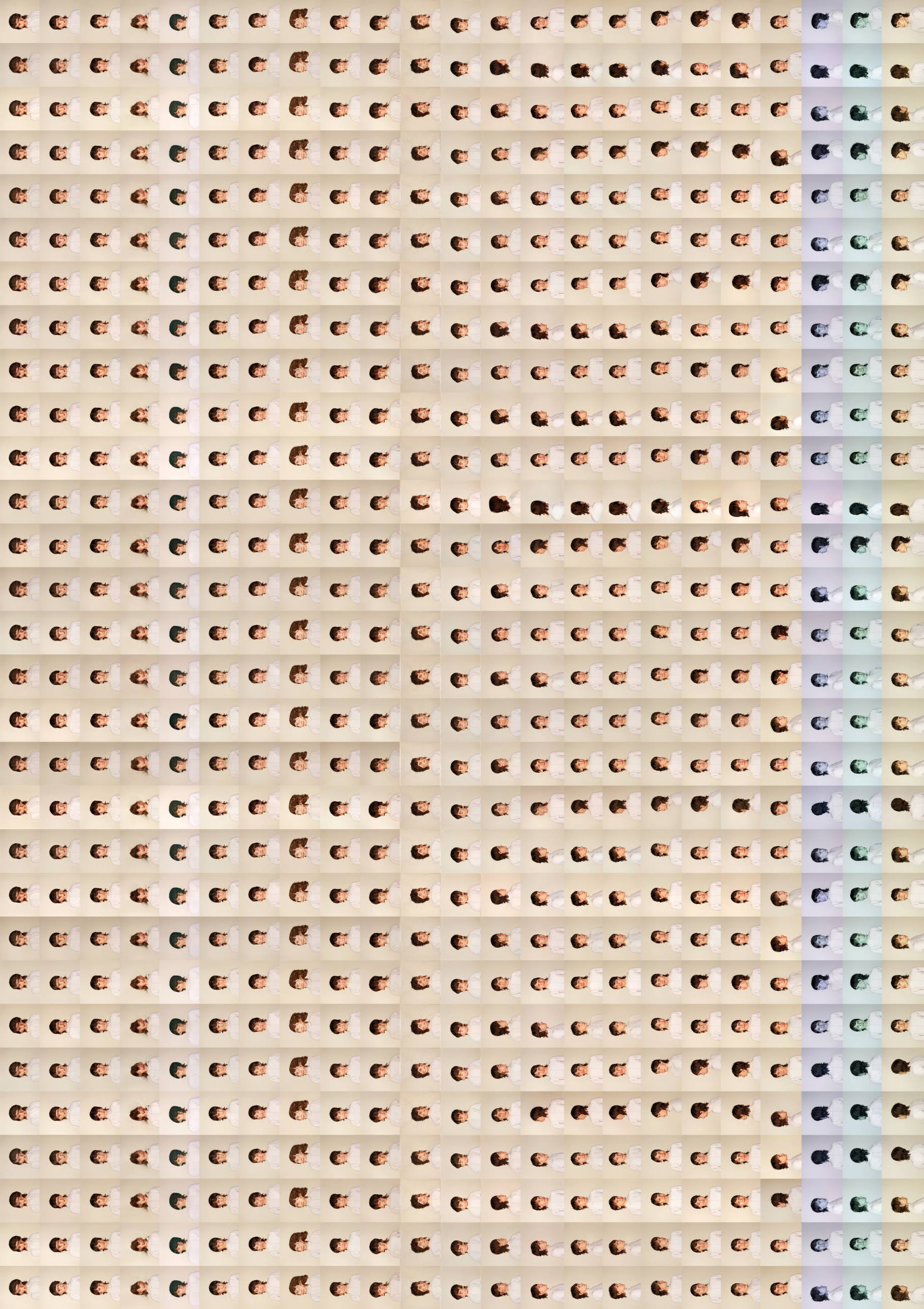
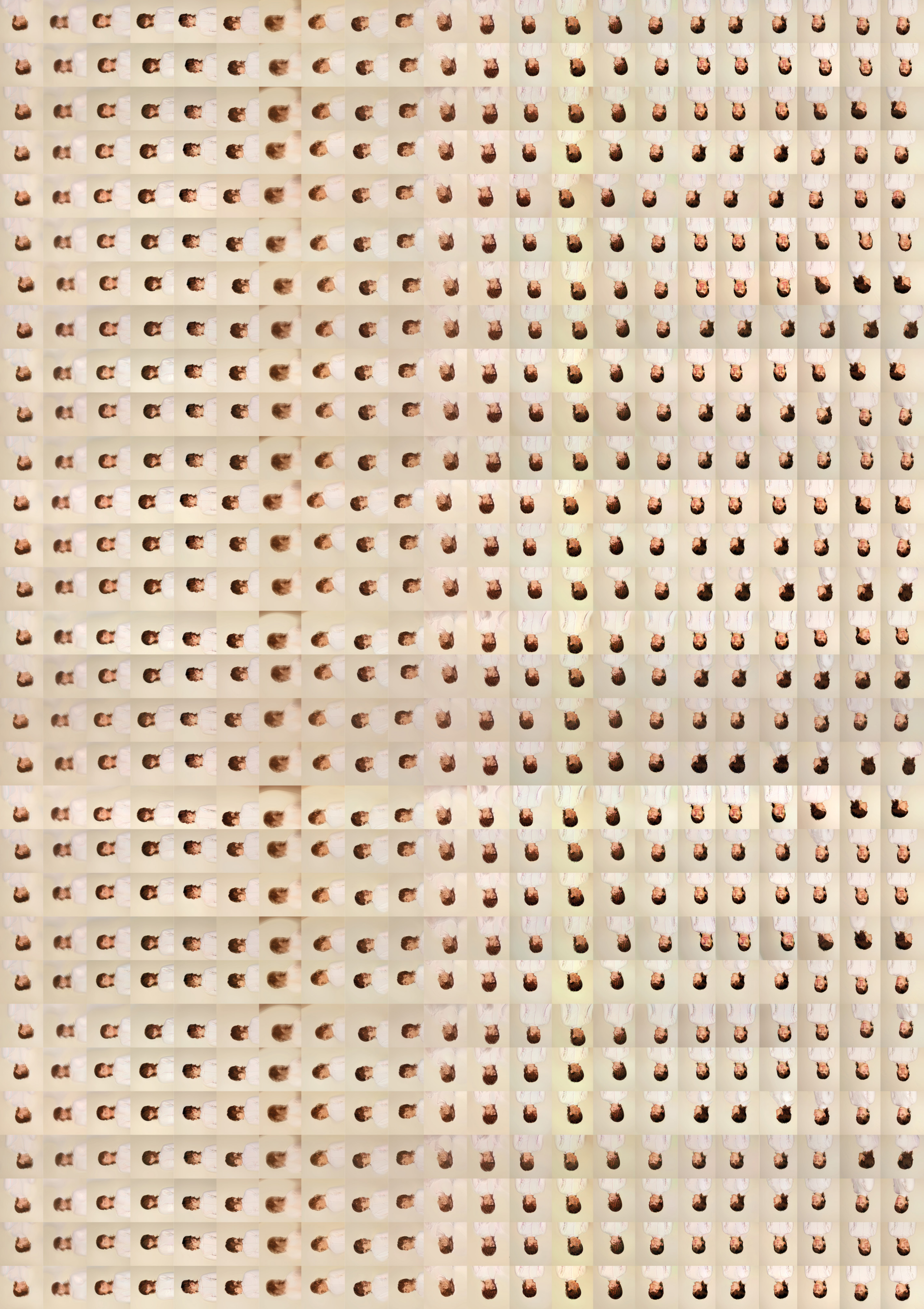


FIG 23 – Training StyleGAN3

training parameters: --cfg stylegan3-t --batch 4 --gamma 0.2 --gpus 1



THE TOOLBOX

As established in earlier chapters, the goal was to choose Machine Learning methods to be used in a creative fashion as an extension to our speculative concept. Intrigued by the notion of translating textual concepts directly into visual representations, as OpenAI first announced it in conjunction with their framework DALL·E[76], we searched for openly accessible Machine Learning frameworks that would allow us to experiment with this procedure. DALL·E was introduced by OpenAI, publishing the paper "Zero-Shot Text-to-Image Generation" as a framework capable of generating an image from written text descriptions. The most notorious example of how well this new architecture performs was an image of an armchair in the shape of an Avocado, created with the same language prompt. OpenAI released CLIP[77], the part of the code that processes language and evaluates how well an image fits a specific text description. While standard image models jointly train an image feature extractor and a linear classifier to predict some label, CLIP consists of an image encoder and a text encoder and was trained on vast and random web scraped images and their respective captions. Instead of predicting the exact caption of an image, it predicts the most appropriate in contrast to all the captions learned. With its capability of predicting how well an image fits a textual description, CLIP can be utilized for guiding a search through the latent space of a given Generative Adversarial Network to find latents that map to the textual concept of images. Since OpenAI published its research on DALL·E in the beginning of 2021 alongside a smaller version of CLIP, researchers and artists have tried to reverse engineer DALL·E and utilize CLIP in conjunction with various Generative Adversarial Networks. The artist and mathematician Katherine Crowson @RiversHaveWings[78] published a Google Colab Notebook that combines CLIP with VQGAN. The latter is short for "Vector Quantized Generative Adversarial Network" and was proposed in the paper "Taming Transformers"[79] at Heidelberg University (2020). VQGAN combines convolutional Neural Networks (traditionally used for images) with Transformers (traditionally used for language).

The combination of the two models allows artists to create various exciting visuals merely by inputting text. This can be a caption, a poem, a lyric or a word. Katherine Crowson found out that she can use CLIP to guide VQGAN towards an image that best matches a given text. CLIP functions as the Perceptor and VQGAN as (Image-)Generator. CLIP is used to guide a search through VQGAN's latent space to find images that match a text prompt very well. Analogous to our concept, we find this process simulates a constant dialogue, as the artist is always prompting, evaluating, and re-prompting the machine. This might represent a symbiotic, interspecies relationship.

In the following, there will be insights into the workflow of creating the short film "Borrowed Limbs" and some best practice examples. It will mainly concentrate on using PyTTI5 beta, a Google Colab notebook developed by sportsracer48[80] and based on the notebook mentioned above by Katherine Crowson. Google Colab notebooks are a hosted Jupyter notebook service. The notebooks are interactive web applications that allow for running Python code directly from the browser in combination with explanatory text and multimedia resources. Google Colab notebooks run on a virtual machine in the Google cloud. Once initiated, the notebook will connect to a Google GPU and/or CPU according to the

76 – "DALL·E: Creating Images from Text," OpenAI, January 5, 2021, https://openai.com/blog/dall-e/.

77 – "CLIP: Connecting Text and Images," OpenAI, January 5, 2021, https://openai.com/blog/clip/

78 – "(16) Rivers Have Wings (@RiversHaveWings) / Twitter," Twitter, accessed January 27, 2022, https://twitter.com/RiversHaveWings.

79 – Patrick Esser, Robin Rombach, and Björn Ommer, "Taming Transformers for High-Resolution Image Synthesis," ArXiv:2012.09841 [Cs], June 23, 2021, http://arxiv.org/abs/2012.09841.

80 – "Sportsracer48 Is Creating AI Art Systems," Patreon, accessed January 27, 2022, https://www.patreon.com/sportsracer48.

requirements of the notebook. Optionally it is always possible to utilize one's own hardware. The fact that Google Colab notebooks are stored and integrated with Google drive and are very easy to access and share made them very popular among researchers. This is especially true for the Machine Learning field as most of the tasks rely very heavily on GPUs and very powerful ones. The option to connect to a Google GPU makes it possible for anyone to experiment with python code, regardless of the specifications and configurations of their own machine.

Google offers the Google Colab pro tier, which gives guaranteed access to their more powerful GPUs (usually a P-100 compared to a K4 without a pro tier) for a monthly subscription of 9.99$ per month. For most of the scenes of our video, we used a PyTTi5-beta notebook as it allows for video input and other added functionalities. This enhanced version of Katherine Crowson's Colab notebook is only available for patrons of sportsracer48 for a tier of 5$/month. Patreon is a membership platform that allows artists and creators to regularly receive a self-deter/mined amount of money from their fans or people using their products. We also used a myriad of other Colab notebooks built upon the Katherine Crowsons original notebook and were developed and released in Patreon's Discord by other members of the PyTTi community. "Recovery DiscoDiffusion v4 Preview Alpha w/ Zooming and Chigozienri keyframes" was released by the user Somnai[81] and allowed for using the SLIP model by Microsoft instead of CLIP and has an implementation of an advanced cutout method implemented by a different user as well as added animation techniques such as Diffusion zooming and Chigozie keyframing. While Somnai is the maintainer of the Disco Diffusion notebooks, many of the improvements made in these notebooks are based on the efforts of other developers, who are all referenced by Somnai.

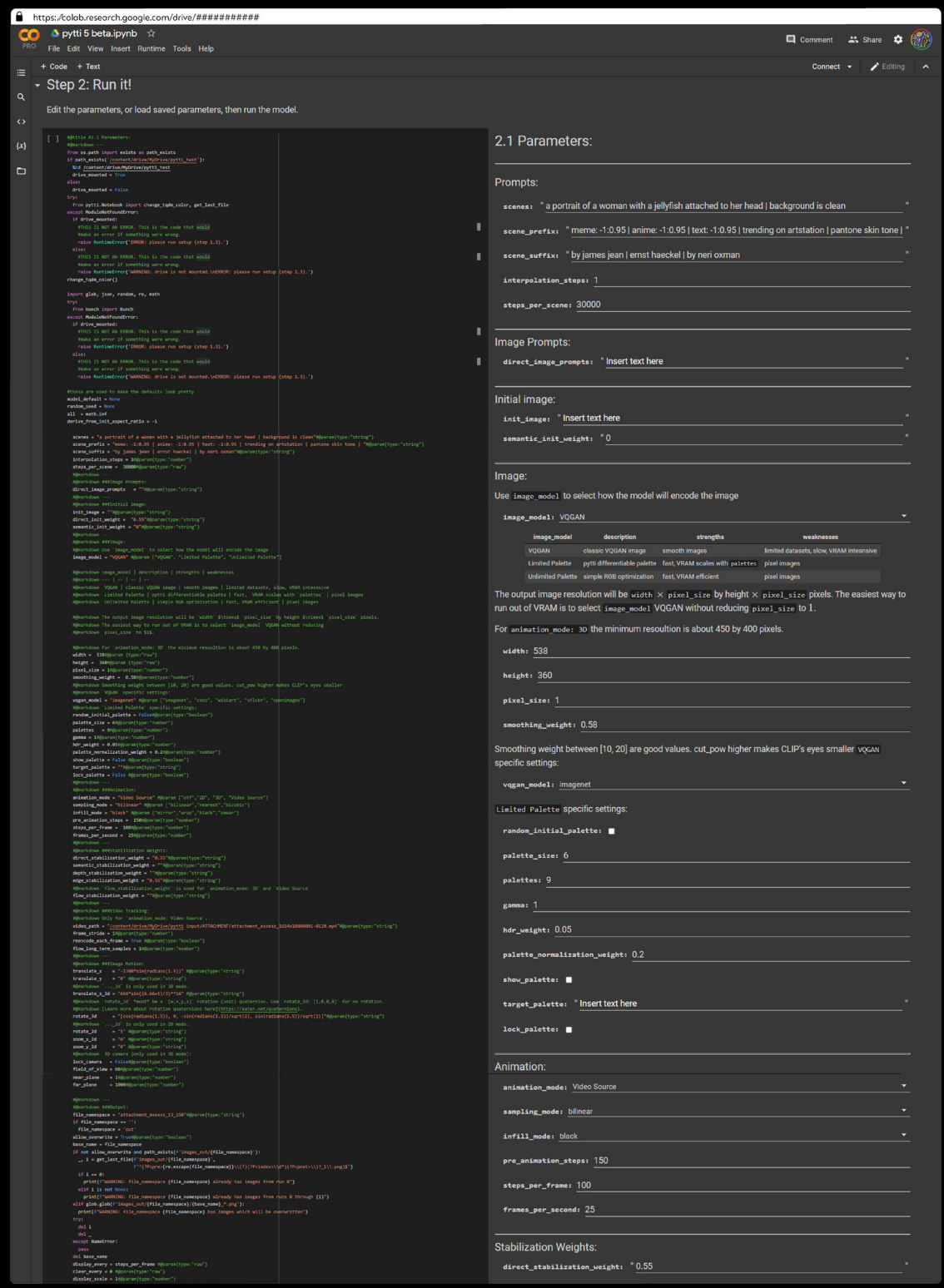


FIG 24 – PyTTi5-beta

81 – "(16) SOMNAI (@Somnai_dreams) / Twitter," Twitter, accessed January 27, 2022, https://twitter.com/Somnai_dreams.

BEING IN A LOOP OF DIALOGUE

At this point, image synthesis with Machine Learning tools is still a time-consuming process and requires quite some devotement to render a short animation with a PyTTi5-beta notebook and the like – as described in detail in the Chapter „The Toolbox“[82]. Nevertheless, some learnings were made while working on this thesis that can speed up the process and improve the results.

The key to appealing image creation is the prompt text design. There can be a limitless combination of prompts, whereas one prompt is limited to 77 tokens of around 38 words. While writing and combining them, one should keep in mind that CLIP uses natural language processing. This means a prompt must not be a string of keywords but rather form coherent sentences. Therefore, it is helpful to have a concept in mind while formulating them. The more distant the prompts are, the bigger the chances that the outcome will be devolving into an abstract mix of shapes and textures. In the beginning, it is best to get started with a single prompt and see how that affects the result. The prompts are separated by a pipe symbol (|). When working with more than one prompt, it is helpful to think about conceptual overlaps.

```
text_prompt = [„A portrait of a woman | The woman is wearing a jellyfish on her head
| Cnidaria“]
```

The prompt “head” refers to a body part of the generated woman, influenced by the first prompt. This will specify the placement of the second prompt. The third prompt, “Cnidaria “, the species the jellyfish belongs to, will open up more variation in the found imagery for the generated headdress. That way a clear connection is recognizable, and the composition and place of the generated fragments will naturally fall into place. Instead of painting with colors it is like painting with concepts. When concepts are designed with care, it is possible to stack more prompts together and still receive a coherent result instead of a random combination of image artifacts.

FIG 25 – Generated Headdress

82 – Chapter „The Toolbox“, p. 66

To better understand the text/image pairs on which CLIP was trained, it is beneficial to test prompts in a search engine to get a rough idea of the imagery activated in the chosen network. Of good use is the clip-retrieval search tool[83] since it is more representative than, for example, a Google search result. With the CLIP retrieval tool it is possible to predict the pictures that will guide the render for a given prompt and hence understand the visual outcome to a certain degree. It also helps to add more adjectives and specifiers to restrict the search. Specifiers could be "seen from a Distance", "seen from above", "imposing", "macro", "close-up" or "portrait". CLIP for example also understands different civilizations and time periods very well, so prompts like "futuristic" or "from the 70s" will have a particular effect.

There is only a finite number of pixels in an image, which means there cannot be infinite details. Too many prompts might compete for the limited space. Therefore, it is important to balance them with other parameters to assure a consistent image outcome. It is effective to use adjectives or specifiers of scale to specify small and large details. That way, it can be defined how many of a particular object should appear in an image.

```
text_prompt = [„A fruit bowl on a table | fruits:5_minions“]
```

Thus, the weight "minions" would only apply to parts of the sythesized image that match the *meaning* fruit, thus turning the fruits into minions. The additional digit scales the number of minion fruits to five.

FIG 26 – Bowl of Minions

83 – "Clip Front," accessed January 27, 2022, https://rom1504.github.io/clip-retrieval/?back=https%3A%2F%2Fknn.laion.ai&index=laion_400m_128G&query=paintings+by+Frida+Kahlo&useMclip=false.

In the same way, there is the possibility to work with negative and positive weights. Some prompts, for example, can influence the overall color, like "martian landscape". The imagery of planet Mars is primarily shades of red. If the sky should still be blue in the generated image outcome, a negative weight could help.

```
text_prompt = [„blue sky:10 | martian landscape | red sky:-1“]
```

In this example, the "blue sky" prompt is additionally amplified by factor 10. Weights can especially come in handy when discovering text fragments like signatures or logos in the outcome imagery. By excluding those images with negative weights, the results will become cleaner. Since a prompt with a negative weight will often go haywire without a stop and erase itself entirely from the final composition, it is better to specify a range.

```
text_prompt = [„blue sky:10 | martian landscape | red sky:-1:-.9 | text:-1:-.9“]
```

A stop should be between **0** and **1** for positive prompts or between **–1** and **0** for negative prompts. Lower stop values will have more effect on the image.

FIG 27 – Blue Sky on Mars

To specify a specific artistic style, there is an option to extend the prompt with an artists' name.

```
text_prompt = [„An Astronaut | sculpture by Yayoi Kusama“]
```

The figure **(FIG 28)** is a collection of images with the same parameters paired with different artist names. Depending on the artists' degree of popularity, it can be necessary to extend the prompt by specifying the artistic specialty. For example, "Sculpture by Yayoi Kusama" or "Paintings by Frida Kahlo". One should be aware that distinctive stylistic features or motifs will be integrated into the output imagery. This can be seen with the Yayoi Kusama prompt and the integration of her trademark, the polka dots. Colored dots that she paints on canvases, sculptures, and people. Frida Kahlo's self-portraits were created throughout all creative phases in her life. The reason she gave herself: Loneliness and the fact that she knows herself best.[84] Thus the generated images often show her face and have not only a stylistic influence. These examples indicate that it is useful to gather knowledge on various art techniques in photography, painting, architecture, or famous artists to incorporate them. Furthermore, it is advantageous to have a basic understanding of harmonic image compositions as that will also influence the quality of the outcome.

FIG 28 – Artist style comparison

84 – "Ein zentrales Thema sind Frida Kahlos Selbstportraits," Singulart Magazin, September 29, 2020, accessed January 27, 2022, https://blog.singulart.com/de/2020/09/29/frida-kahlos-selbstportraits/.

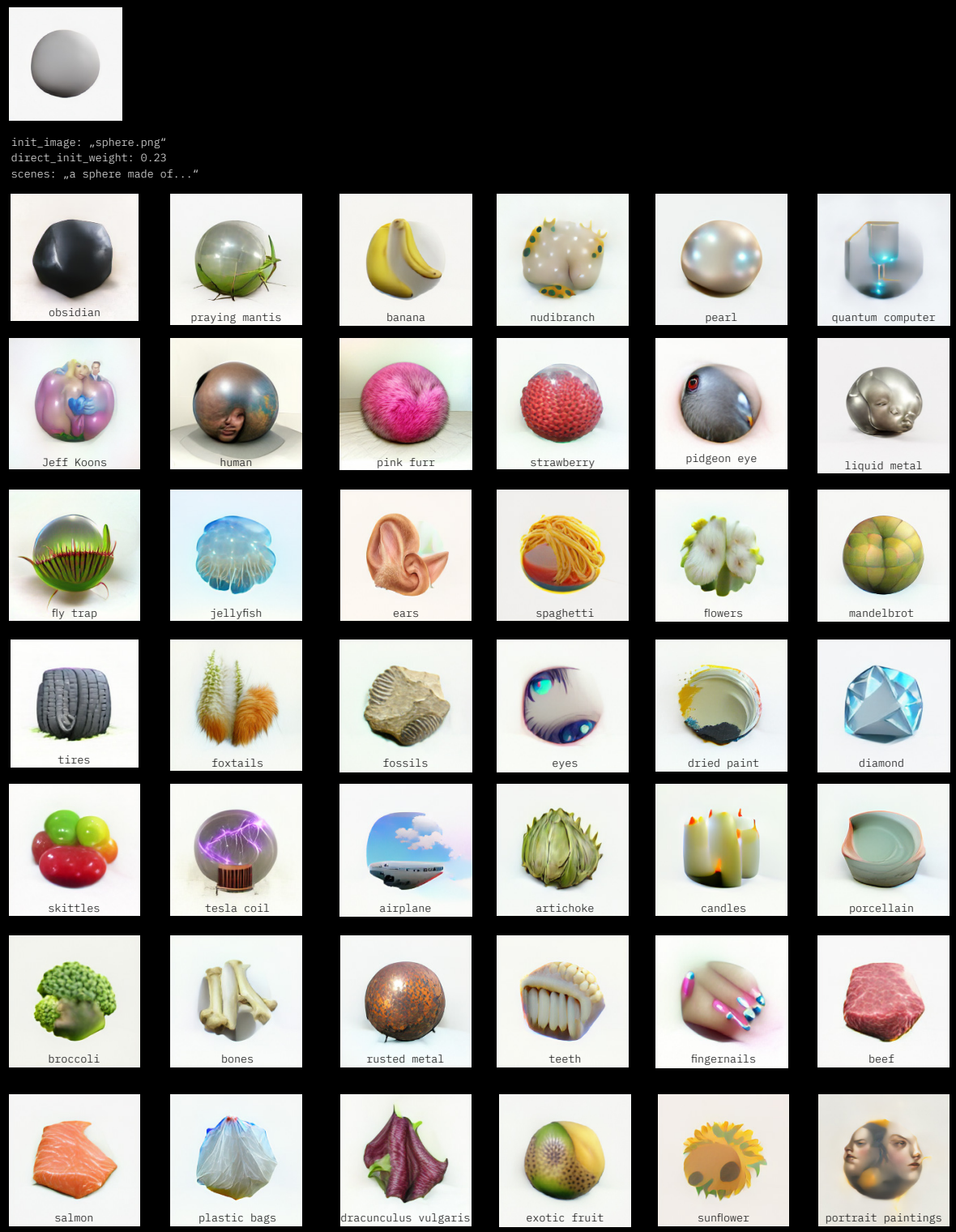


FIG 29 – Texture and material studies

Instead of artists, it is, of course, also possible to use games or filmmakers with prominent stylistic features, such as the game “No Man’s Sky” to integrate the color schemes or “Image composition by Wes Anderson” to incorporate the symmetrical film settings. Keywords often used on art platforms or social media can also be influential. “Unreal engine”, “Trending on Artstation” and “Macro art” are just some examples. Also, a specific art medium can serve as an interesting stylistic prompt like “oil on canvas”, “woven textile”, “ceramic sculpture” or, to give some digital examples, “Blender”, “Octane Renderer” and “Houdini”.

While thinking about the colors of a generated image, it can be helpful to incorporate objects and materials with that color into the prompt design to achieve the desired look. A simple one would be “banana” to use yellow while balancing the distinctive, curved shape with negative weights. An example for the opposite could be the wish to use a texture “snake” but leveling out a common color coming with this prompt: “green:-1:-.9”. Some material and texture examples can be seen in figure **(FIG 29)**.

The comparisons shown in the figures **(FIG 28-29)** were created with an initial image prompt. As one of the additional parameters, there can be an image defined that will be used as overall guidance for the composition and overall color scheme. The shaded sphere used in the examples was held in grey shades on purpose to influence the composition mainly. It ensures the central motif is generated inside the sphere. That way, the main composition stays mostly the same to ensure the comparability of different artist and texture prompts.

On the other hand, an initial image can intentionally influence the colors, which makes it always helpful to use a good result as an initial input in the next iteration — or to create an image with the desired color scheme

beforehand. This approach is more efficient in achieving a coherent look and, for this reason, applied in our production process. An initial image can also be sketches to define an image composition. In figure **(FIG 30)**, a simple sketch was used to influence the position of "Frida Kahlo as an Astronaut". Applying even rudimental hints of facial features, they are picked up and applied to the output. Using the sketch additionally in the parameter "direct_image_prompts" in combination with a positive "direct_init_weight" functions like a mask where the generation only takes part inside the sketched shapes. The "direct_init_weight" parameter can be used to determine the influence of the initial image on the final composition. This is relevant while working with several prompts to define a hierarchy. Values between **[1, 10]** can be used depending on the purpose.

Next to the prompt and initial image parameters, the PyTTi5-beta notebook offers over 60 additional parameters, which help specify and polish the resulting image. The most influential for the described process and the creation of "Borrowed Limbs" will be discussed in the following.

As already mentioned in the previous chapter[85], sportsracer48 enhanced the original notebook by Katherine Crowson with the option for video input among other features and functionalities. A video can be given as input, and each frame will get generated with the set prompts. Depending on the defined output image resolution and assigned GPU by Google Colab or the local computer, this process can take several hours. Therefore, choosing either short sequences, lower resolutions, or fewer frames to be generated makes sense. We opted for the latter two options during the process but developed a workaround to compensate for the loss of quality. This will be described in more detail in the next chapter[86].
Input videos can be actual footage filmed with a camera and 3D renders. In the figure on the previous page are some stills of a rendered video in Cinema4D with an octane render engine. The video shows an Avatar falling back into nothingness. The given prompt was the all-time favorite "Frida Kahlo as an Astronaut". **(FIG. 31)**

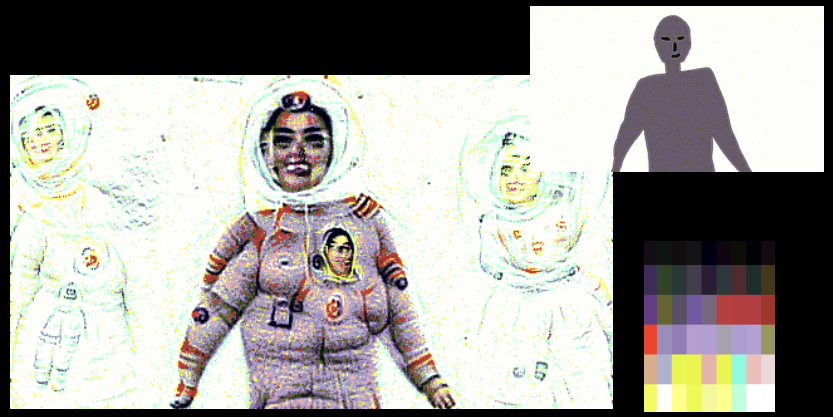
A parameter which indicates the number of steps between each generated frame is "steps_per_frame". The more steps the further developed is a frame. While experimenting to achieve a desired prompt design concept, it makes sense to set this value lower in combination with a small output resolution to see the influences of the prompt design quicker, for example, **50**. The prompts will be less influential if the value is lower than **10**. When happy with the output, the value could be set to **250** or even up to **500** to receive more details.

"infill_mode" specifies how newly generated pixels should be filled if they appear on edge. "Mirror" reflects the image over the boundary while "wrap" pulls them from the opposite side. "Black" will simply color them black. The "smear" input samples the closest pixel around.

Stabilization weights are helpful values to hierarchize the inserted inputs. The "direct_stabilization_weight" will function as a "direct_image_prompt" for video frames. It uses the current frame as input. The higher the value, the closer the result to the initial video frame.
„semantic_stabilization_weight" and "edge_stabilization_weight" keep the contours of the image consistent from frame to frame in case of a video input. "flow_stabilization_weight" prevents flickering in video sources.

85 – Chapter „The Toolbox", p. 66

86 – Chapter „Designing a Pipeline for Machine Collaboration", p. 92



Limited color palette **image_model**

Rough Silhouette with facial features as **init_image** and **direct_image_prompt**

Rough Silhouette as **init_image**

FIG 30 – Sketch input

3D video as **init_image** input
no **stabilization_weights** result in visually inconsistent frames

FIG 31 – 3D rendering input

The parameter “smoothing_weight“ keeps the color palette smaller and closer to another. Values around **`[1, 6]`** will give detail and contrast, while higher values result in a desaturated look. With a value of **75**, the colors are getting close to monochromatic. A well-working default value lies around **`[10, 20]`**, balancing details and a clear motif.

Another parameter influencing the images’ contrast is “gamma”. But it should be used with care since values below **`0.8`** in combination with a “hdr_weight” higher than **`0.25`** result in very uniform images. Default values for “gamma” are **`[1, 1.5]`**, for “hdr_weight” **`[0.01, 0.25]`**.

The “Cut_pow” parameter influences the number of details in an image. Using values around **`[0.2, 0.6]`** output larger shapes useful while setting up an overall scene composition, for example, a person in a full-body shot. Is the image more about the fine details, “Cut_pow” values around **`[0.6, 1.7]`** can give a close-up shot of only the face much more wanted details. In general, it can be used as a scale and proportion specifier for the given prompts. The value influences how CLIP “sees” the image. In a visualization of Remi Durant[87], shown as a screenshot in **(FIG 32)**, it can be seen how CLIP cuts the image into smaller segments. These smaller segments are scored for accuracy. According to Durant, a natural bias seems to happen around the center of an image in the way the cut regions are picked. This leads to CLIP putting more effort in generating the center of an image but leaves the edges with less coverage.

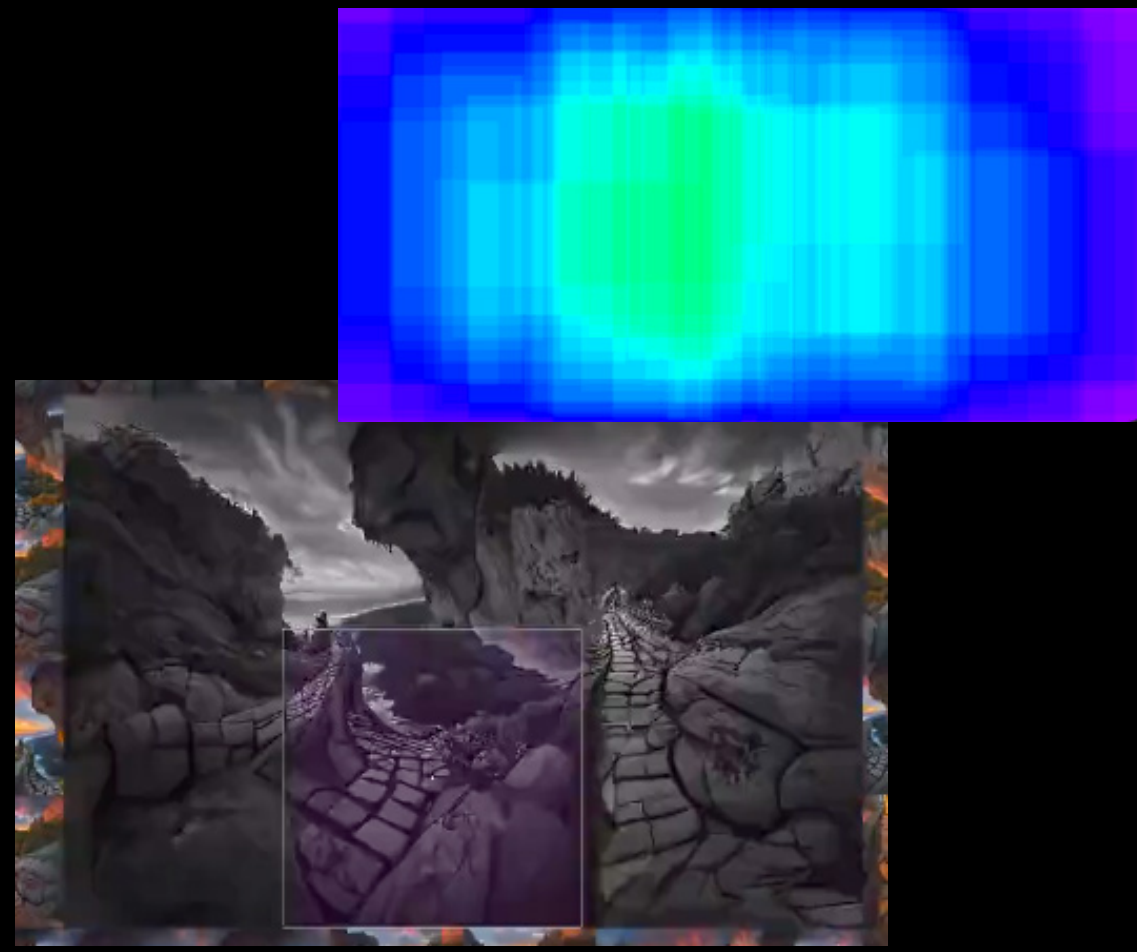


FIG 32 – How CLIP „sees“ the image

87 – @remi_durant, accessed January 27, 2022, https://twitter.com/remi_durant/status/1460607677801897990

While testing and figuring out the best working prompts, a shared table of keywords was created on Notion[88]. On the one hand, this was part of the thesis's documentation and a source of inspiration for new and unusual prompt designs. An efficient workflow also saves well-working parameter combinations as JSON files with an image displaying the expected result. The PyTTi5-beta notebook includes an extra cell that can be used for copying and pasting the JSON to make use of earlier iterations as well setting up batch execution, which is of course limited by the Google Colab maximum runtime for the pro tier which is capped at 24 hours.

scene04-parameter.json


```
{
  "scenes": "A female human head in profile | her hair is made of snakes",
  "scene_prefix": "Trending on Artstation | ",
  "scene_suffix": "Frida Kahlo | Surrealism",
  "interpolation_steps": 160,
  "steps_per_scene": 60100,
  "direct_image_prompts": "",
  "init_image": "",
  "direct_init_weight": "",
  "semantic_init_weight": "",
  "image_model": "VQGAN",
  "width": 640,
  "height": 360,
  "pixel_size": 4,
  "smoothing_weight": 5,
  "vqgan_model": "imagenet",
  "random_initial_palette": true,
  "palette_size": 6,
  "palettes": 9,
  "gamma": 1,
  "hdr_weight": 0.15,
  "palette_normalization_weight": 0.2,
  "show_palette": true,
  "target_palette": "",
  "lock_palette": false,
  "animation_mode": "off",
  "sampling_mode": "bilinear",
  "infill_mode": "smear",
  "pre_animation_steps": 250,
  "steps_per_frame": 50,
  "frames_per_second": 2,
  "direct_stabilization_weight": "",
  "semantic_stabilization_weight": "",
  "depth_stabilization_weight": "",
  "edge_stabilization_weight": "",
  "flow_stabilization_weight": "",
  "video_path": "",
  "frame_stride": 1,
  "reencode_each_frame": true,
  "flow_long_term_samples": 1,
  "translate_x": "-1700*sin(radians(1.5))",
  "translate_y": "0",
  "translate_z_3d": "(50+10*t)*sin(t/10*pi)**2",
  "rotate_3d": "[cos(radians(1.5)), 0, -sin(radians(1.5))/sqrt(2), sin(radians(1.5))/sqrt(2)]",
  "rotate_2d": "5",
  "zoom_x_2d": "0",
  "zoom_y_2d": "0",
  "lock_camera": true,
  "field_of_view": 60,
  "near_plane": 1,
  "far_plane": 10000,
  "file_namespace": "default",
  "allow_overwrite": false,
  "display_every": 50,
  "clear_every": 0,
  "display_scale": 1,
  "save_every": 50,
  "backups": 5,
  "show_graphs": false,
  "approximate_vram_usage": false,
  "ViTB32": true,
  "ViTB16": false,
  "RN50": false,
  "RN50x4": false,
  "learning_rate": null,
  "reset_lr_each_frame": true,
  "seed": 13140087269443264454,
  "cutouts": 40,
  "cut_pow": 2,
  "cutout_border": 0.25,
  "border_mode": "smear"
}
```


FIG 34 – Prompt sharing

88 – https://www.notion.so/

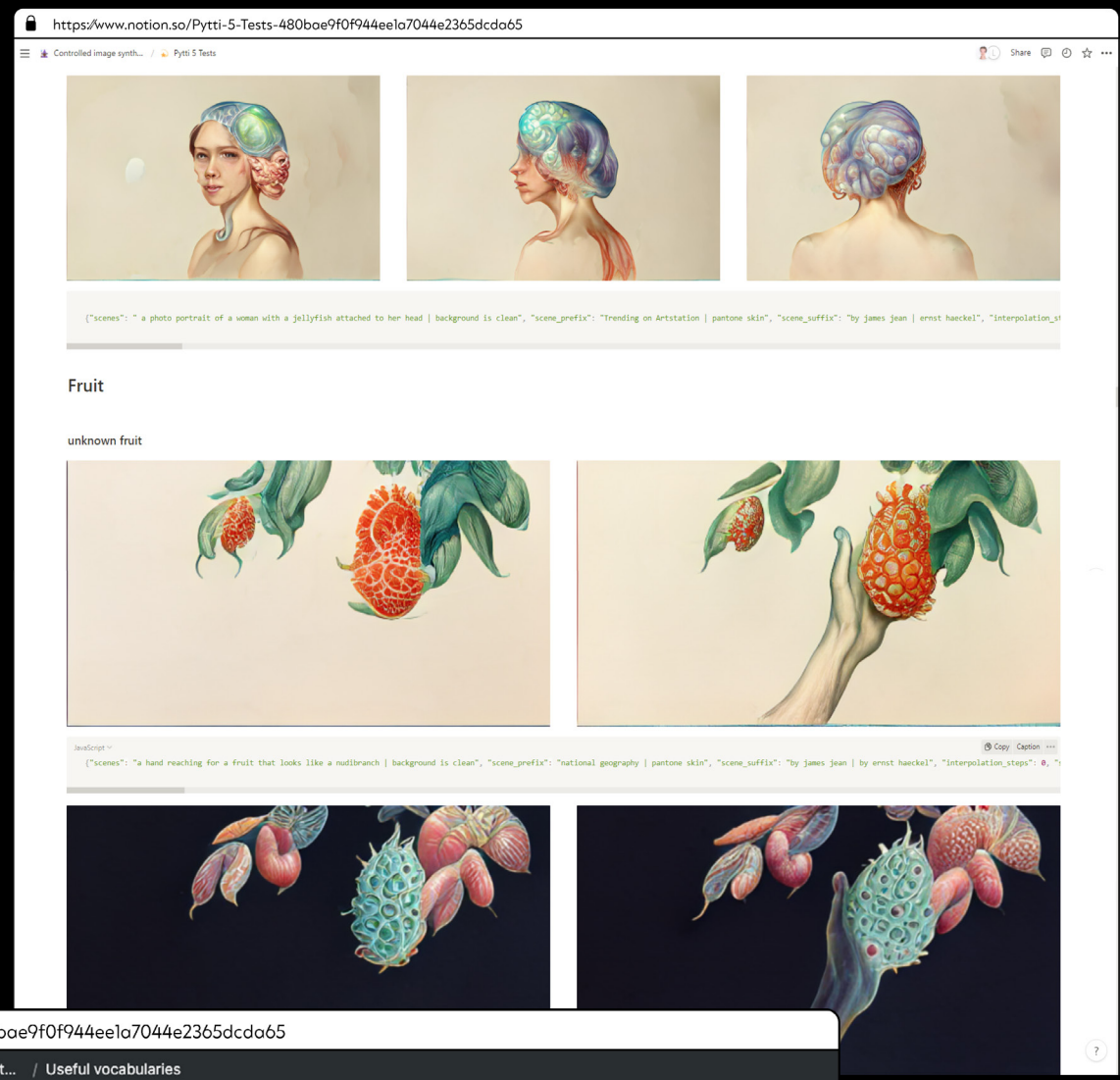


FIG 33 – Prompt Collection

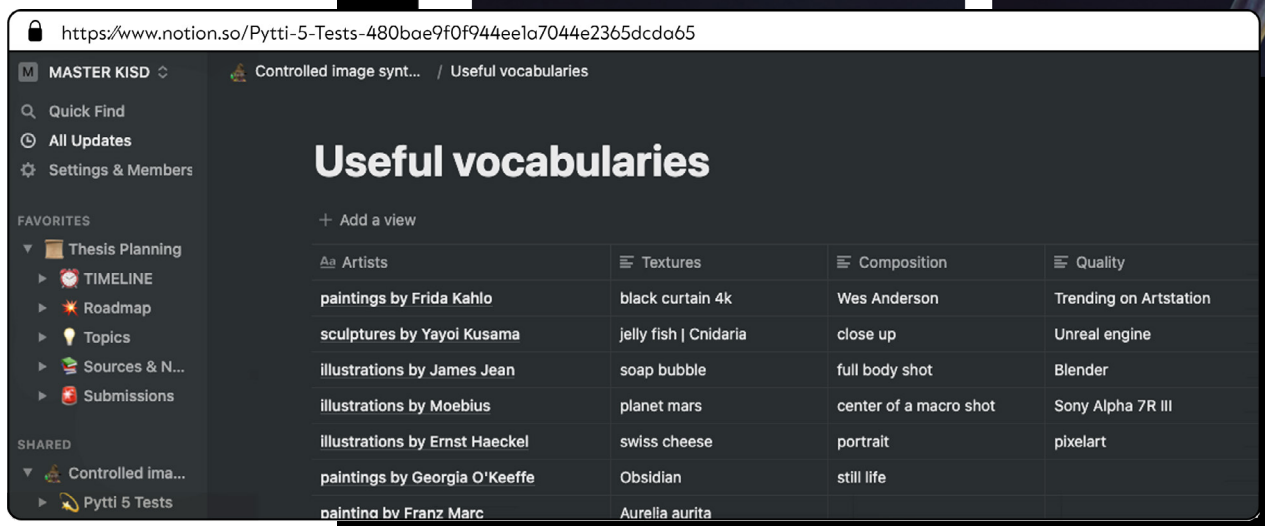


FIG 35 – Vocabularies

FIG 36 – Attachment

text_prompt = [„A portrait of a woman with a nudibranch attached to her head | background is clean“, „scene_prefix“: „trending on artstation | pantone skin tone | „, „scene_suffix“: „by james jean | ernst haeckel | by neri oxman“]

pytti 5 beta.ipynb

FIG 37 – Picking Fruit

text_prompt = [„a hand reaching for a fruit that looks like a nudibranch | background is clean“, „scene_prefix“: „national geography | pantone skin“, „scene_suffix“: „by james jean | by ernst haeckel“

pytti 5 beta.ipynb

For the short film “Borrowed Limbs” one of the main goals was to generate a cohesive, cinematic narration in the aesthetic. Image results are very dependent on the initial image input. Prompt designs that have worked with other input images might not work with the current ones and must be adjusted. The following figures show prompt design and essential parameters combined with the input image. We focused on the overall composition and color scheme while creating the initial footage. Important shapes and image composition were created with objects or color contrasts. The AI-human attachment shape and position, highlighted in our speculative film, was arranged on the head of the model with a bathing cap and two woolen balls in different colors. To make the end result look more grown to the head, we faded the color of the skin tone to match the color of the bathing cap. With the desire to create an unusual fruit, we altered one woolen ball with colored pin needles and created a composition with cut leaves from a bush. More input and generated impressions can be seen on the next pages. **(FIG 36-40)**

An overview of style frames was created to specify the scenes’ look and integrate the prompt design into the overall concept. It also shows several stages of connection between the AI and its human device. The stages are detachment, symbiotic fusion, attachment and inner view. Each stage is accompanied by specifically designed prompts and adjusted parameters for the input footage to define a visual language.
(FIG 41)

FIG 38 – A warm hug

text_prompt = [„two people are hugging | their bodies are melting together in a bright light explosion | background is clean“, „scene_prefix“: „Trending on Artstation | pantone skin |“, „scene_suffix“: „ | Magic Realism Of Rob Gonsalves :1 | paintings by Jim Burns:2 | illustrations by ernst haeckel:1“]

pytti 5 beta.ipynb

FIG 39 – Symbiosis

text_prompt = [„A portrait of a woman with a nudibranch attached to her head | background is clean“, „scene_prefix“: „national geography | pantone skin“, „scene_suffix“: „by james jean | by ernst haeckel“]

pytti 5 beta.ipynb

FIG 40 – Her hands

text_prompt = [„Hands are turning | each finger looks like a sensory device | LED lights are blinking at the end of each finger“, „scene_prefix“: „painting by james jean:.5 | illustration by ernst haeckel:1 | Trending on Artstation„]

pytti 5 beta.ipynb

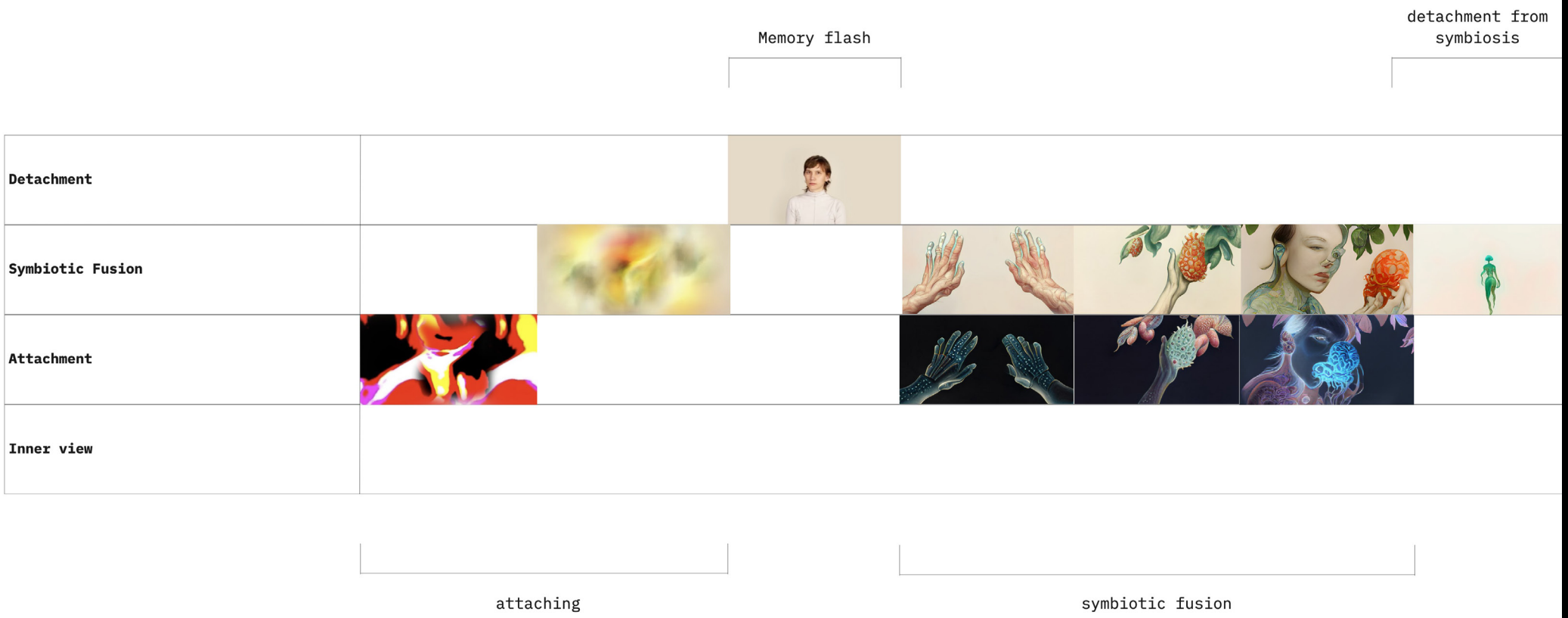
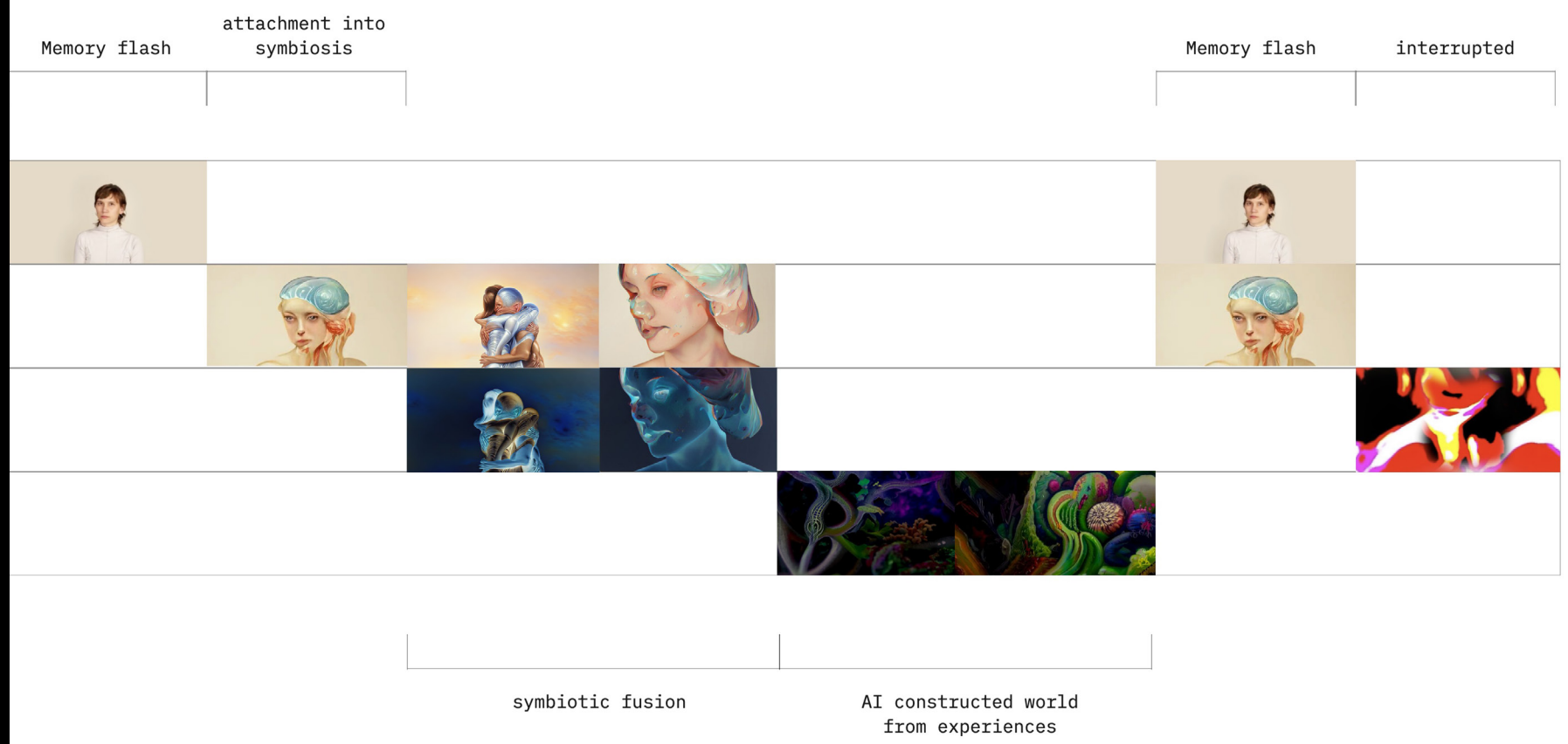


FIG 41 – Styleframes



DESIGNING A PIPELINE FOR MACHINE COLLABORATION

Our production pipeline is a confluence from the learnings of the last chapter and solutions that had to be found for mediating limitations imposed by a lack of time and compute power. Synthesizing images with image diffusion techniques extensively explained in the last chapter, combined with inference results from the custom trained StyleGAN3 model, resulted in imagery with resolutions and framerates uncommon for traditional video footage production. The image diffusion process performed with the help of PyTTi5-beta, and other previously mentioned Colab notebooks, processes videos frame by frame. As a very high image quality and clear details are very important for the diffusion model to catch on the details when prompted with an image, motion blur typically occurring in video frames is not ideal. Videos with framerates of 25+ frames per second would also take an excessive time for processing for a few seconds of video. Therefore, it was decided to use stop motion photography and interpolate between the frames with the help of another Neural Network. For the interpolation between our stop motion frames, we used the RIFE Framework. RIFE - Real-Time Intermediate Flow Estimation for Video Frame Interpolation[89] supports arbitrary-timestep interpolation between a pair of images and provides automatic video processing ability. With RIFE, a video rendered from stop motion frames can be processed, and the motion gaps between the images will be filled with inferred images. This tool works exceptionally well for reconstructing dynamic changes (e.g., motion and color between images) and was included in the processing pipeline proposed in this thesis. As the processing time of image generation with diffusion models will increase exponentially with the output image's resolution, it was decided to downscale the stop motion video and 3D renderings. The images were processed with PyTTi5-beta at a resolution of 860x360 - four times smaller than the resolution aimed at for the final cut: 3440x1440. After testing a myriad of neural-network-backed upscaling tools like ESRGAN and RealSR in conjunction with different models found in the Upscale Wiki Upscaling database,[90] the best solution for our needs turned out to be Real-ESRGAN.[91] The successor of ESRGAN was developed at the Shenzhen Institutes of Advanced Technology of the Chinese Academy of Sciences in conjunction with Shanghai AI Laboratory by Xintao Wang et al.

Quite handily, this implementation comes with a portable version that can be executed via the command line, with the only parameters that need to be specified being the in- and output directory and the upscaling Neural Network to be used. The Neural Network "realesrganx4plus" included in the GitHub code repository worked best with the use case of producing the video for this thesis work. Other Neural Networks that are more suitable for upscaling 2D cartoons are included in the repository as they appear to have been trained on Anime footage.

The video footage was also created with a custom trained StyleGAN3[92] model as described in detail in the Chapter "A Latent Space full of emotions".[93] To create the dataset, ~2000 portrait shots of the same human photo model with different head poses and with varying expressions were made at a photo studio. The resulting images were automatically cropped and exported in the *.jpg format, using the Software Capture One[94]. The StyleGAN3 repository comes with a dataset preparation tool that zips the images and distributes them evenly into folders for optimized

89 – hzwer, RIFE - Real-Time Intermediate Flow Estimation for Video Frame Interpolation, Python, 2022, https://github.com/hzwer/arXiv2020-RIFE.

90 – "Model Database - Upscale Wiki," accessed January 28, 2022, https://upscale.wiki/wiki/Model_Database.

91 – Xintao, Real-ESRGAN, Python, 2022, accessed Jnuary 30, 2022, https://github.com/xinntao/Real-ESRGAN.

92 – NVlabs/Stylegan3, Python (2021; repr., NVIDIA Research Projects, 2022), https://github.com/NVlabs/stylegan3.

93 – Chapter „A Latent Space full of emotions", p. 50

94 – "Capture One Pro Photo Editing Software," accessed January 28, 2022, https://www.captureone.com/en.

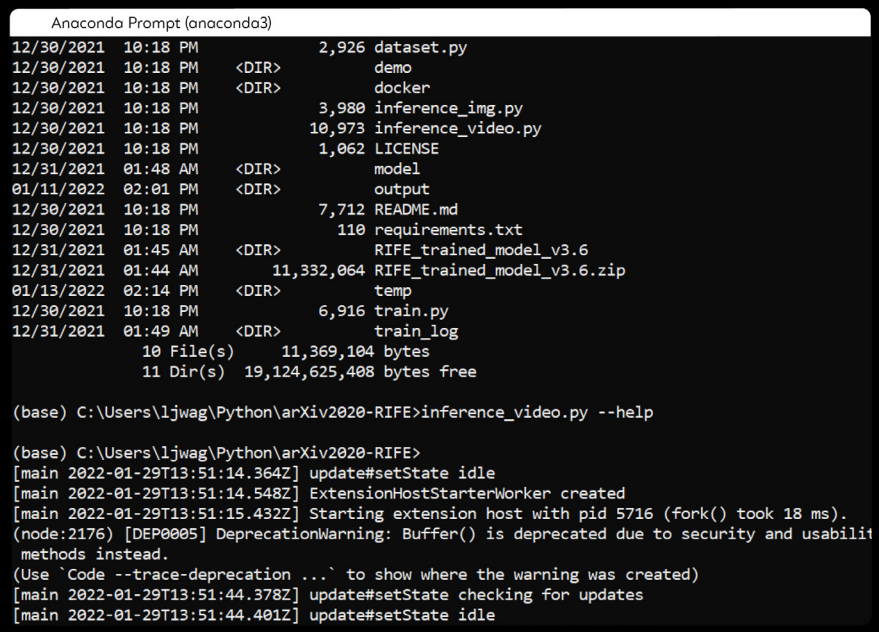


FIG 42 – Frame interpolation

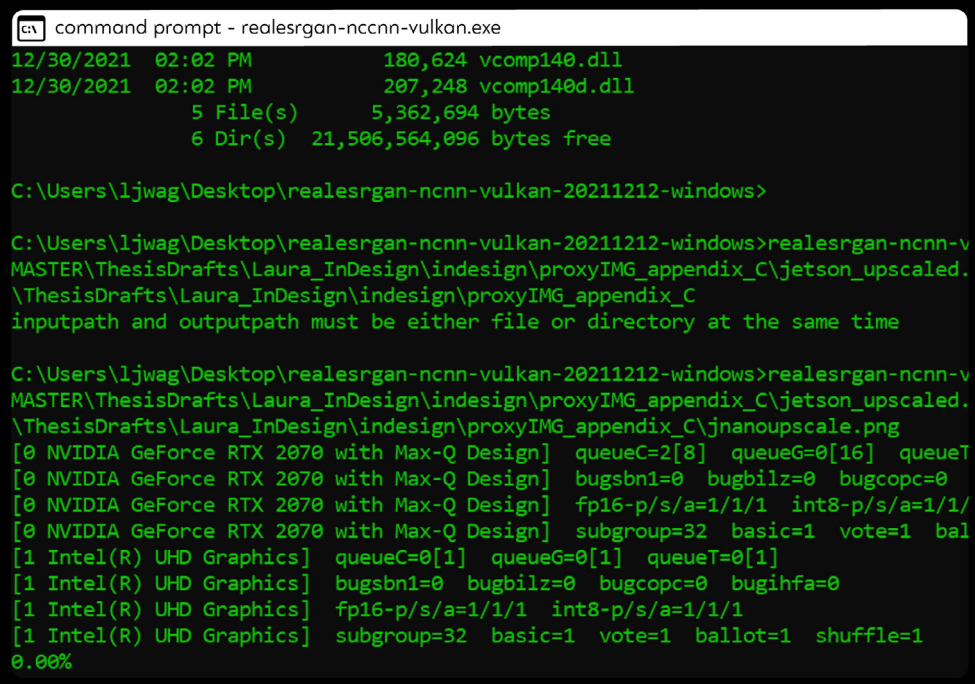


FIG 43 – Upscaling

batch processing. The training process is started with the command line, specifying different arguments, e.g. input and output directories, batch size, number of GPUs to be used, and intervals in which model files are saved. During training, visual renderings of the Neural Networks learned representations are outputted periodically. This enables monitoring of the training process. The decision to restart the training with different parameters can be made based on the observed results. If the training process is interrupted, the training of the Neural Network file can be resumed with the same data. Another option is to resume the training in a process called transfer training or fine-tuning – this is usually done with other or modified datasets. The Neural Network files resulting from the training are saved in the *.pkl file format. The Python Pickle is a compressed format that allows for Python objects (for example tensor objects) to be serialized on disks and deserialized (loaded into the program memory), a process referred to as "pickling" and "unpickling".[95] Image synthetization or the generation of images with a trained StyleGAN model is generally referred to as inference.

The next step of the proposed pipeline is to generate a couple of images from different seeds. Seeds are points in the latent space, the abstract space of possible visual variations. The best images are picked on purely visual curational criteria from the seed images. These seeds can then be used to perform other inference operations to get the desired visual outcome. Latent space interpolations are interpolations between different seeds that denote places in the latent space. This process is also called "latent space walk" or "sightseeding" as it is the visualization of gradually moving from one place to another in the latent space of the Neural Network. Another process used during the exploration of techni-

95 – "PKL File Extension - What Is a .Pkl File and How Do I Open It?," accessed January 28, 2022, https://file-info.com/extension/pkl.

ques in this thesis is called truncation-interpolation. The Machine Learning enthusiast and educator Derrick Schulz built some additional tools for StyleGAN2-ada-Pytorch[96] and released them on GitHub, allowing for additional operations done with a trained Neural Network.
One of these operations is called truncation-interpolation. Truncation-interpolation allows for choosing a seed and diverging from its origin in latent space away from the center of the latent space and saving this process in a video. As mentioned in the chapter "A Latent Space full of emotions"[97] the bigger the truncation value, the more the generating algorithm will "wander off" into the latent space, where the less possible noisy results are located. Truncation-interpolation videos show the process of the visual representation of a seed 'dissolving' into noise. Diego Porres released some additional functionalities on top of the StyleGAN3 implementation by NVlabs in his own GitHub repository: StyleGAN3-fun. One of the added scripts is called "discriminator dreaming" and allows for visualizing the neuron's weights of the discriminator algorithm instead of the generator's. A visualization of this process can be seen in the previous chapter.[98]

To this date, inference from StyleGAN models always results in square resolutions. As the hardware configurations of the computer used for training the StyleGAN model used in this thesis only allowed for training models with an output resolution of 256x256, we used the upscaling procedure elaborated earlier in this chapter with these images and videos as well. The audio recordings for the video were made in a recording studio using a Neumann TLM 102 Microphone in conjunction with Adobe Audition. The sound files were later used to assemble a dataset for training a WaveGAN[99] model. A WaveGAN is the same type of Neural Network as StyleGAN, but for sounds. When trained on human speech samples, WaveGAN models can generate sounds that resemble human speech.
For the post-processing, color grading, type animations, and the final cut, we used the AdobeCC Software and Blender.

This thesis proposes a modular pipeline for controlled image synthesis with AI tools, as visualized in the following diagram. **(Fig 44)**

96 – Derrick Schultz (he/him), Dvschultz/Stylegan2-Ada-Pytorch, Jupyter Notebook, 2022, https://github.com/dvschultz/stylegan2-ada-pytorch.

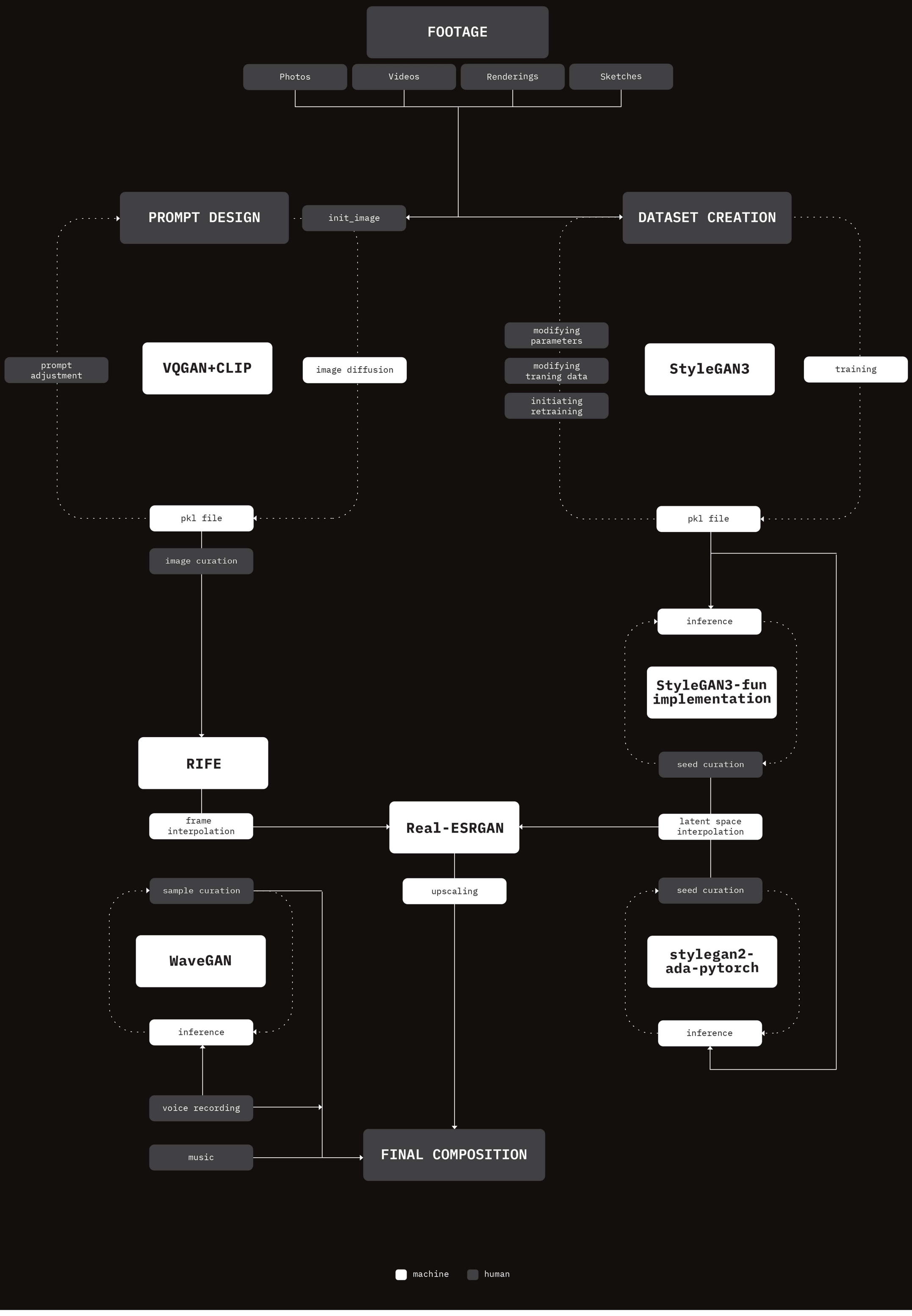
97 – Chapter „A Latent Space full of emotions", p. 56

98 – FIG 18, p. 53

99 – Chris Donahue, WaveGAN (V2), Python, 2022, https://github.com/chrisdonahue/wavegan.

MACHINE COLLABORATION

PRODUCTION PIPELINE



CONCLUSION

The future visions by posthumanist thinkers described at the beginning of this thesis set the direction of further research and resulted in the development of a speculative concept. However, neither does the concept turn away from the species human nor does it understand the "post" prefix as the end of humanity. Instead, it is intended to represent alternatives to humanistic views that focus on a particular concept of a ***human*** being. Technological developments such as AI should not be understood as a transhumanistic extension but as an independent species that needs the context of the material world just as humans do. Derived from approaches in cognitive science and philosophy of embodiment and interspecies companionship, we focus on a vision of a symbiotic relationship of the organic and inorganic. Thinking is related to a body. This fact makes it crucial to practice interrelationships and value a symbiotic coexistence. The speculative concept focuses on a future AI that utilizes the human body to gain an embodied understanding of its environment. Transhumanist concepts in which the ***human*** body is technologically enhanced are appropriated and inverted, and the human body itself becomes an enhancement in the role of a sensing device. Connected in a constant dialogue, organism and machine are mutually dependent. *Meaning is synthesized* through collaboration and embodied context.

These notions are reflected in the collaborative design method developed throughout this thesis. Video footage is drastically reinterpreted and altered with AI techniques in conjunction with natural language prompts. The practical exploration of appropriating and creating with AI as an "unruly actor" yielded insights into the potential impact of disruptive technologies on the design profession. Repetitive tasks always have been automated with technology – but with AI, designers are gaining a tool that can be used to co-create in a drastically new way. A promising approach would be to view AI as a collaborator rather than a competitor, a colleague of a new species.

Designers will be able to outsource tedious tasks and engage with machines collaboratively, generating and curating from an infinite latent space of possible designs. As a machine whisperer, the designer will enter a constant dialogue in setting the goals, parameters, and constraints before reviewing and curating the results. Concluding, adjusting, and repeating a generative process before the result can be manually refined will become an important task. AI enables many industries to move away from the prerogative of a one size fits all standardization approach for their products. A new approach of radical personalization is emergent. This could open up the possibility for designers to develop entirely new types of experiences. Realtime AI rendering in an upcoming generation of video games, prompt brushes in software like Photoshop, AI texture packs – these developments will change the way designers work. But while looking forward to these developments, one should also take a critical look at the foundation these technologies are built on.

Exploring AI tools in a designerly way led to an examination of one component of AI frameworks from which behaviors result: Training data. A closer look at the dataset used for training can explain why visual representations created by an AI consistently evoke the same idealized and sexualized body features and poses when describing female bodies. Working on visualizations of a female body solely with text prompts was difficult. One had to be careful with adjectives or the naming of particular body parts to avoid insinuating images. This clearly showed one of

the problems of working with Machine Learning methods. Most datasets are web scraped - programmatically copied from the internet, and thus they inherit and carry forward biases from their birthplace. This dilemma is exacerbated by the fact that AI is introducing artificial content into the vast corpus of our collective representation of online knowledge, the internet. Regurgitated representations of past data might forever haunt us, as they forever get entangled with the original material of the present: a vicious cycle.

When appropriating tools, a designer should critically reflect on the technology they are working with. The main goal in creating the short film "*Borrowed Limbs*" was, first, to make sure that the narrative of the film addressed the problems that AI technology brings. Second, to design a production pipeline for machine collaboration with the short film as a best practice outcome for image synthetization. While most of the tools were released to the public less than a year ago, the workflow was based on experimentation and flexibility. Prompting the machine, evaluating the result, and refining the following prompt to get the anticipated result – a human-machine dialogue. The goal was to *synthesize meaning* through a coherent narration. The proposed production pipeline serves as a prototype for future machine-human collaboration. It allowed us to research through our chosen design method how a possible interaction could proceed.

Since the image generated by the machine is not necessarily the one the human has in mind, this workflow requires practice. But with knowing the vocabulary and the proper emphasis, it is like having learned a new language. This language opens up to a new way of creating visualizations for speculative concepts. Sketches, raw photography, quick 3D renderings, and the correct settings can be combined into a coherent, visual language. These experimental approaches might seem rough in their origin and describe a workflow far off the glossy and polished approaches often seen in futuristic visuals. They testify to the joy of experimentation and use of fundamental design principles of shapes, textures, and composition combined with the appropriation of new technologies.

THANK YOU

During this thesis, we had the great help of **Bessie Normand**, who lent us her face and let us copy it into the latent space of our human sensing device. **Mayssa Kaddoura**, who owns the voice of an advanced AI, and convinced us that she is fully embodied.

FIG 45 – Her world

BIBLIOGRAPHY

„ASUNDER - TEGABRAIN." http://tegabrain.com/Asunder. Accessed October 7, 2021.

„BIOMEDIEN|ZKM." Accessed January 21, 2022. https://zkm.de/en/exhibition/2021/12/biomedien.

BRAIDOTTI, ROSI. "A Theoretical Framework for the Critical Posthumanities," Theory, Culture & Society 36, no. 6 (November 1, 2019): 31–61, Accessed January 22, 2022, https://doi.org/10.1177/0263276418771486.

BRAIDOTTI, ROSI. The interview with Judith Butler (S. W. Oh, Trans. Korea). Journal of Feminist Theories & Practices (1999).

BRAIDOTTI, ROSI, AND HLAVAJOVA, MARIA. Posthuman Glossary. New York; London: Bloomsbury Publishing Plc, 2018.

"CLIP: CONNECTING TEXT AND IMAGES." OpenAI, January 5, 2021, Accessed January 28, 2022, https://openai.com/blog/clip/.

CROGAN, PATRICK. "Bernard Stiegler: Philosophy, Technics, and Activism." Cultural Politics 6, no. 2 (July 1, 2010): 133–56. https://doi.org/10.2752/175174310X12672016548162.

„CRYONICS INSTITUTE." Accessed January 22, 2022. https://www.cryonics.org/.
Dagar, Sumit. "Design in the Age of AI." Medium, April 18, 2020. https://uxdesign.cc/design-in-the-age-of-ai-aa0385e82efc.

DAGAR, SUMIT. "Design in the Age of AI," Medium, April 18, 2020, Accessed 18 Jan. 2022, https://uxdesign.cc/design-in-the-age-of-ai-aa0385e82efc.

"DALL·E: CREATING IMAGES FROM TEXT." OpenAI, January 5, 2021, Accessed January 28, 2022, https://openai.com/blog/dall-e/.

„DATAISM." In Wikipedia, January 10, 2022. https://en.wikipedia.org/w/index.php?title=Dataism&oldid=1064746251.

DUNNE, ANTHONY AND RABY, FIONA. „Speculative Everything: Design, Fiction, and Social Dreaming", (Cambridge, Massachusetts; London: The MIT Press, 2013).

ESSER, PATRICK, ROMABCH, ROBIN AND OMMER, BJÖRN. "Taming Transformers for High-Resolution Image Synthesis," ArXiv:2012.09841 [Cs], June 23, 2021, Accessed January 28, 2022, http://arxiv.org/abs/2012.09841.

EVELETH, ROSE. "Bodyhackers are all around you, they're called women", Accessed 16 Jan. 2022, https://splinternews.com/bodyhackers-are-all-around-you-they-re-called-women-1793856408.

GANE, NICHOLAS. "When We Have Never Been Human, What Is to Be Done?: Interview with Donna Haraway." Theory, Culture & Society 23, no. 7–8 (December 1, 2006): 135–58. https://doi.org/10.1177/0263276406069228.

„GOOGLE." "Understanding Searches Better than Ever Before," October 25, 2019. https://blog.google/products/search/search-language-understanding-bert/.

GU, YEON JEONG. "The Disembodiment of Digital Subjects and the Disappearance of Women in the Representations of Cyborg, Artificial Intelligence, and Posthuman." Asian Women 36, no. 4 (December 31, 2020): 23–44. https://doi.org/10.14431/aw.2020.12.36.4.23.

„HABITAT_MP3D DATASET." Accessed January 22, 2022. https://aihabitat.org/datasets/hm3d/.

HARARI, YUVAL NOAH. „21 Lessons for the 21st Century", London: Penguin Random House UK 2018.

HAYLES, N. KATHERINE. How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics. Chicago, Ill: University of Chicago Press, 1999.

HOCHSCHUH & DONOVAN. "Empathy Swarm." Accessed January 21, 2022. https://hochschuh-donovan.com/portfolio/empathy-swarm/.

HUI, JONATHAN. "GAN — Why It Is so Hard to Train Generative Adversarial Networks!," Medium (blog), October 29, 2019, Accessed January 28, 2022, https://jonathan-hui.medium.com/gan-why-it-is-so-hard-to-train-generative-advisory-networks-819a86b3750b.

HZWER. RIFE - Real-Time Intermediate Flow Estimation for Video Frame Interpolation, Python, 2022, https://github.com/hzwer/arXiv2020-RIFE.

„IBM 701." In Wikipedia, December 1, 2021. Accessed 18 Jan. 2022, https://en.wikipedia.org/w/index.php?title=IBM_701&oldid=1058120081.

KÄRKKÄINEN, KIMMO, AND JUNGSEOCK JOO. "FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age." ArXiv:1908.04913 [Cs], August 13, 2019. http://arxiv.org/abs/1908.04913.

MELANIE MITCHELL. Artificial Intelligence_ A Guide for Thinking Humans (2019, Farrar, Straus and Giroux) - Libgen.Lc.Epub," n.d.

„NVIDIA." „Knight Rider Rides a GAN with AI, NVIDIA Omniverse." 16 Apr. 2021, Accessed 18 Jan. 2022, https:/blogs.nvidia.com/blog/2021/04/16/gan-research-knight-rider-ai-omniverse/.

„NVIDIA." „NVIDIA Announces Omniverse Open Beta, Letting Designers ...." 5 Oct. 2020, Accessed 18 Jan. 2022, https:/nvidianews.nvidia.com/news/nvidia-announces-omniverse-open-beta-letting-designers-collaborate-in-real-time-from-home-or-around-the-world.

„NVIDIA." „Omniverse Audio2Face AI Powered Application | NVIDIA." Accessed 18 Jan. 2022, https:/www.nvidia.com/en-us/omniverse/apps/audio2face/. Accessed 18 Jan. 2022.

PETRIČ, ŠPELA. „PL'AI", Accessed January 21, 2022. https:/www.spelapetric.org/#/plai/.

"PKL FILE EXTENSION - WHAT IS A .PKL FILE AND HOW DO I OPEN IT?." accessed January 28, 2022, https:/fileinfo.com/extension/pkl.

PORRES, DIEGO. "Discriminator Synthesis: On Reusing the Other Half of Generative Adversarial Networks," ArXiv:2111.02175 [Cs, Eess], November 12, 2021, http:/arxiv.org/abs/2111.02175.

"PUBLIC RECORD OF ENRON MAIL AND EMAIL ARCHIVE." Accessed October 7, 2021. http:/www.enron-mail.com/.

RADFORD, ALEC, JONG WOOK KIM, CHRIS HALLACY, ADITYA RAMESH, GABRIEL GOH, SANDHINI AGARWAL, GIRISH SASTRY, ET AL. "Learning Transferable Visual Models From Natural Language Supervision." ArXiv:2103.00020 [Cs], February 26, 2021. http:/arxiv.org/abs/2103.00020.

RIBEIRO, GUIDA. "Posthumanmonsters.", Accessed January 21, 2022. https:/posthumanmonsters.net/about.

ROMERO, ALBERTO. "GPT-4 Will Have 100 Trillion Parameters — 500x the Size of GPT-3." Medium, September 11, 2021. https:/towardsdatascience.com/gpt-4-will-have-100-trillion-parameters-500x-the-size-of-gpt-3-582b98d82253.

SAMPLE, IAN, AND IAN SAMPLE SCIENCE EDITOR. "Why Is the Internet so Overwhelmingly Male?" The Guardian, October 19, 2018, sec. Technology. https:/www.theguardian.com/technology/2018/oct/19/global-inequalities-make-the-internet-overwhelmingly-male.

SAVVA, MANOLIS, ABHISHEK KADIAN, OLEKSANDR MAKSYMETS, YILI ZHAO, ERIK WIJMANS, BHAVANA JAIN, JULIAN STRAUB, ET AL. "Habitat: A Platform for Embodied AI Research." In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 9338–46. Seoul, Korea (South): IEEE, 2019. https:/doi.org/10.1109/ICCV.2019.00943.

SCHMITHUESEN, BIRK. "SpeculativeAI Series The SpeculativeAI Series Consists of Aesthetic Experiments Designed to Make Processes of Artificial Neural Networks Perceptible to Humans through Audiovisual Translation.," accessed January 22, 2022, http:/www.birkschmithuesen.com/_SpeculativeAI.

SHAPIRO, LAWRENCE AND SPAULDINGO, SHANNON. „Embodied Cognition", The Stanford Encyclopedia of Philosophy (Winter 2021 Edition), Edward N. Zalta (ed.) Accessed January 22, 2022, https:/plato.stanford.edu/archives/win2021/entries/embodied-cognition.

SIMS, ANDREW. "The Problems with PredictionThe Problems with Prediction: The Dark Room Problem and the Scope Dispute: The Dark Room Problem and the Scope Dispute." Philosophy and Predictive Processing, 2017. Accessed January 22, 2022, https:/doi.org/10.15502/9783958573246.

SINGULART MAGAZIN. "Ein zentrales Thema sind Frida Kahlos Selbstportraits," September 29, 2020, Accessed January 28, 2022, https:/blog.singulart.com/de/2020/09/29/frida-kahlos-selbstportraits/.

SOUFLIS, ZOË. "Post-, Nicht- und Parahuman", in „Future Bodies. Zur Visualisierung von Körpern in Science und Fiction", ed. Marie-Luise Angerer, Kathrin Peters und Zoë Soufoulis. 273–300, Wien, New York: Springer-Verlag 2002.

"SPECULATIVEAI SERIES THE SPECULATIVEAI SERIES CONSISTS OF AESTHETIC EXPERIMENTS DESIGNED TO MAKE PROCESSES OF ARTIFICIAL NEURAL NETWORKS PERCEPTIBLE TO HUMANS THROUGH AUDIOVISUAL TRANSLATION." Accessed January 22, 2022. http:/www.birkschmithuesen.com/_SpeculativeAI.

STANGL, WERNER. Online Lexikon für Psychologie und Pädagogik'. Online Lexikon für Psychologie und Pädagogik. https:/lexikon.stangl.eu, Accessed 11.01.2022

"START OF AI - GOOGLE SEARCH." Accessed January 22, 2022, https:/www.google.com/search?q=start+of+ai&oq=start+of+ai&aqs=chrome..69i57.3662j0j4&sourceid=chrome&ie=UTF-8.

"STRELKA INSTITUTE." Accessed January 22, 2022. https:/strelka.com/en/

THE ARTIFICIAL INTELLIGENCE CHANNEL. The Trouble with Bias - NIPS 2017 Keynote - Kate Crawford #NIPS2017, 2017. https://www.youtube.com/watch?v=fMym_BKWQzk.

„THE ECONOMIST." "Design Takes over, Says Paola Antonelli," Accessed October 02, 2021, https://www.economist.com/news/2010/11/22/design-takes-over-says-paola-antonelli.

TURING, A. M. "I.—COMPUTING MACHINERY AND INTELLIGENCE." Mind LIX, no. 236 (October 1, 1950): 433–60. https://doi.org/10.1093/mind/LIX.236.433.

SYNTHESIZING MEANING
– BIBLIOGRAPHY –

TABLE OF ILLUSTRATIONS

FIG 01 – *HUMAN* Own illustration, Generated with StyleGAN2-ada-pytorch

FIG 02 – MISSING PEOPLE Own illustration, Generated with StyleGAN2-ada-pytorch

FIG 03 – FATE OF HUMANITY Own illustration, Generated with Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

FIG 04 – FIFTY-DIMENSIONAL SPACE Own illustration, Generated with Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

FIG 05 – AI HABITIAT "Habitat_MP3D Dataset," accessed January 22, 2022, https://aihabitat.org/datasets/hm3d/.

FIG 06 – ENRON EMAIL Own illustration

FIG 07 – MACHINE CREATOR, HUMAN CURATOR Own illustration, Generated with Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

FIG 08 – EXP. #2 CONVERSATION "SpeculativeAI Series The SpeculativeAI Series Consists of Aesthetic Experiments Designed to Make Processes of Artificial Neural Networks Perceptible to Humans through Audiovisual Translation.," accessed January 22, 2022, http://www.birkschmithuesen.com/_SpeculativeAI.

FIG 09 – EMAPTHY SWARM "BioMedien | ZKM," accessed January 21, 2022, https://zkm.de/en/exhibition/2021/12/biomedien.

FIG 10 – ASUNDER "Asunder - Tegabrain," accessed October 7, 2021, http://tegabrain.com/Asunder.

FIG 11 – POSTHUMAN MONSTERS "Posthumanmonsters," accessed January 21, 2022, https://posthumanmonsters.net/about.

FIG 12 – PL'AI "KERSNIKOVA_Spela_Petric_PLAY_20201206_HanaJosic-69.Jpg (1500×2247)," accessed January 22, 2022, https://images.squarespace-cdn.com/content/v1/5aeca48a506fbe863b23a8b6/1621256631779-VSRSUM0QU79QS84MT47Z/KERSNIKOVA_Spela_Petric_PLAY_20201206_HanaJosic-69.jpg?format=1500w.

FIG 13 – CONSTANT FLUX OF CREATION Own illustration, Generated with Disco Diffusion v3.1 [w/ SLIP & DangoCutn].ipynb

FIG 14 – TRANSMACHNIST CYCLE Own illustration

FIG 15 – ENTOMON HYBRID Own illustration, Generated with StyleGAN2-ada-pytorch

FIG 16 – HUMAN LIMBS SCULPTURE Own illustration, Generated with pytti 5 beta.ipynb

FIG 17 – HUMAN SENSING DEVICE Own illustration, Generated with pytti 5 beta.ipynb

FIG 18 – NOISE OF NOTHINGNESS Own illustration, Generated with StyleGAN2

FIG 19 – IMPROVEMENT LAYER BY LAYER Own illustration, Generated with StyleGAN3

FIG 20 – FAILED FINE TUNING RESULT Own illustration, Generated with StyleGAN3

FIG 21 – DISCRIMINATOR DREAMING Own illustration, Generated with StyleGAN3

FIG 22 – TRAINING STYLEGAN2 Own illustration, Generated with StyleGAN2

FIG 23 – TRAINING STYLEGAN3 Own illustration, Generated with StyleGAN3

FIG 24 – PYTTI5-BETA Screenshot of the private Google Colab Notebook pytti 5 beta.ipynb

FIG 25 – GENERATED HEADDRESS Own illustration, Generated with pytti 5 beta.ipynb

FIG 26 – BOWL OF MINIONS Own illustration, Generated with pytti 5 beta.ipynb

FIG 27 – BLUE SKY ON MARS Own illustration, Generated with pytti 5 beta.ipynb

FIG 28 – ARTIST STYLE COMPARISON Own illustration, Generated with pytti 5 beta.ipynb

FIG 29 – TEXTURE AND MATERIAL STUDIES Own illustration, Generated with pytti 5 beta.ipynb

FIG 30 – SKETCH INPUT Own illustration, Generated with pytti 5 beta.ipynb

FIG 31 – 3D RENDERING INPUT Own illustration, Generated with pytti 5 beta.ipynb

FIG 32 – HOW CLIP „SEES“ THE IMAGE @remi_durant, accessed January 27, 2022, https://twitter.com/remi_durant/status/1460607677801897990

FIG 33 – PROMPT COLLECTION Screenhot of the private Notion.so space

FIG 34 – PROMPT SHARING Screenhot of a JSON file in Visual Studio

FIG 35 – VOCABULARIES Screenhot of the private Notion.so space

FIG 36 – ATTACHMENT Own illustration, Generated with pytti 5 beta.ipynb

FIG 37 – PICKING FRUIT Own illustration, Generated with pytti 5 beta.ipynb

FIG 38 – A WARM HUG Own illustration, Generated with pytti 5 beta.ipynb

FIG 39 – SYMBIOSIS Own illustration, Generated with pytti 5 beta.ipynb

FIG 40 – HER HANDS Own illustration, Generated with pytti 5 beta.ipynb

FIG 41 – STYLEFRAMES Own illustration

FIG 42 – FRAME INTERPOLATION Screenshot Anaconda prompt

FIG 43 – UPSCALING Screenshot Real-ESRGAN

FIG 44 – MACHINE COLLABORATION PRODUCTION PIPELINE Own illustration

FIG 45 – HER WORLD Own illustration, Generated with StyleGAN2

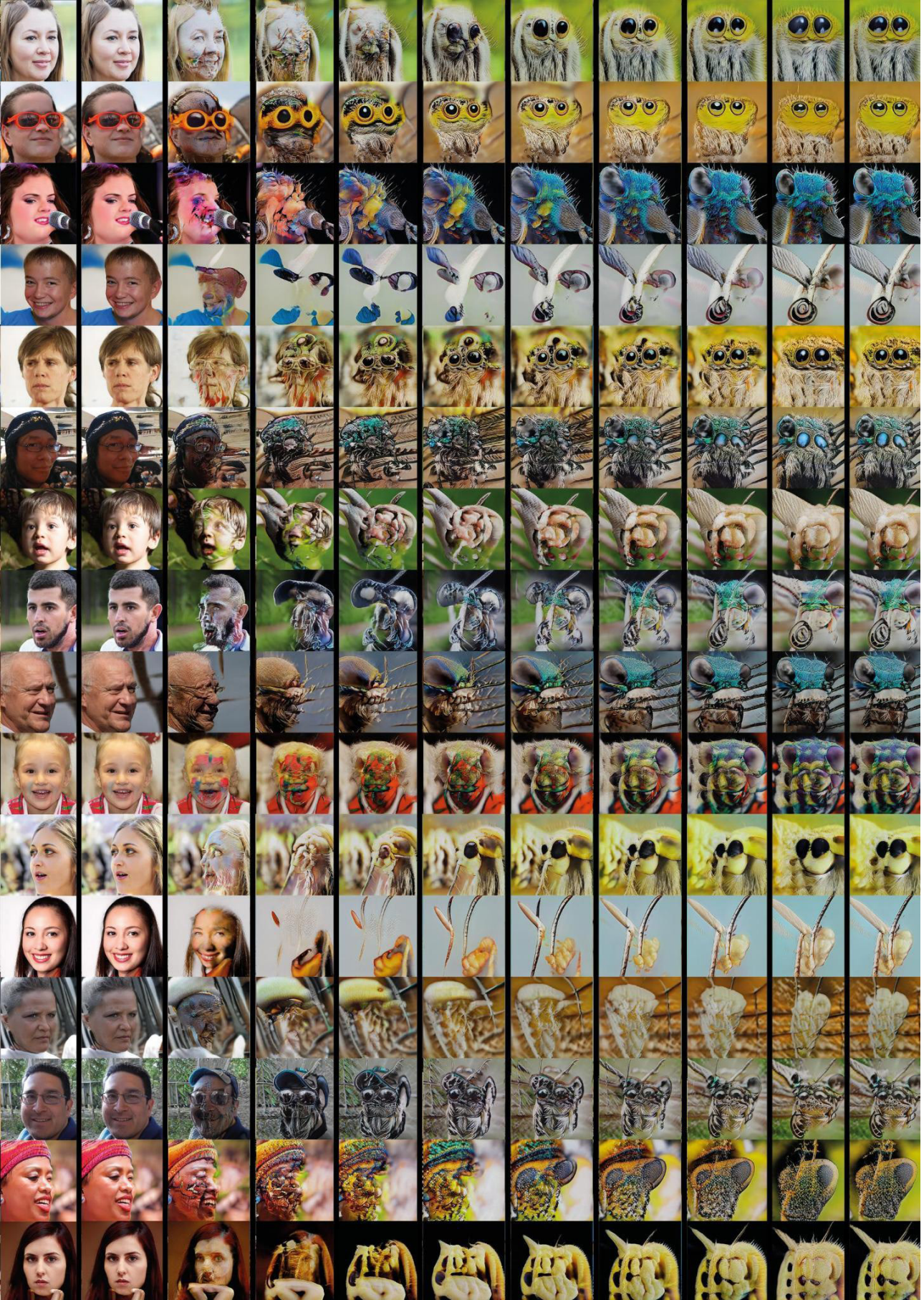
All AI-generated image examples using Google Colab Notebooks in this thesis are combined with the captions indicating the text prompt and other relevant deviations from the default inputs to reproduce the image.

APPENDIX A
INTERSPECIES METAMORPHOSIS

"Entomon – Interspecies metamorphosis" was the result of a self initiated project and resulted in visual experiments speculating on posthuman representation. A Generative Adversarial Network was trained with the StyleGAN2-ada-pytorch framework of NVlabs. We adapted a network that was originally trained on human faces, in transfer training it with our dataset of insect close-ups. The process of training a GAN on two incompatible datasets, produced phenotypes of human-insect hybrids. Insects and arachnids species are seen as disgusting, as vermins which shouldn't be part of the ***human*** habitat. A hybridization between human faces and features of unwanted bodies lets us discover alternatives to the humanist understanding of the „natural". Exploring the latent space of bodily features between species was an inspiring process for us and helped us to think about possible posthuman forms of representation.

In Franz Kafka's novel „Die Verwandlung" the protagonist transforms into a bug in his appearance but remains human in size. This in-between stage alienates him from the world of ***humans*** and the world of bugs. The feeling of non-affiliation is familiar for individuals not fitting the humanist norms of society. But those norms are so narrow, those left out are in a more significant number. Influenced by posthuman visions, our aim is to overcome a binary worldview and embrace hybrid futures and diverse bodies.

In the world of insects, some species undergo metamorphosis in their life. This process is often accompanied by a change of behavior. How much will we need to reshape ***human*** bodies to detach from discriminating norms, from natural and artificial?



TRAINING AND INFERENCE WITH A STYLEGAN NETWORK

We used Stylegan2-ada-pytorch developed by NVlabs (NVIDIA), for training and transfer training Generative Adversarial Networks. Our aim was to explore the whole pipeline (testing pre-trained models - creating a dataset from curated images - training a dataset - generating images from our trained dataset) in order to get familiar with the use of StyleGAN2 as a tool for creative processes. We decided to install stylegan-ada-pytorch on the computers at the university and execute everything from the python console, as for our approach and creative process, it was important to have full control on all the training parameters, automate the generation of videos with batch files and to have a unlimited amount of training attempts.

Unfortunately, the specifications of the computers at the university (NVIDIA RTX 3070 GPU) only allow for training images up to 256x256 pixels. However, we think it is preferable to practice with smaller images, shorter training time and the use of less resources (electricity, time) and later move on to creating dataset with high fidelity images and training them. (when we know what we are doing) While exploring our trained Network we tested different methods such as projection of the latent space to an external image, style mixing - the selective combination of vectors in the latent space, and we rendered linear latent space morph videos. In the future it would be interesting to find more ways to (interactively) explore the network and its latent space.

I INSECT – A POSTHUMAN POEM

Hello blood bag, I'm a vermin, I disgust you - crawling, wiggling, humming, pulsating, secreting secretions - slimy mucous.

Evolution set us apart long ago, and now we converge. I see you already built antennas to attach them to your head, and you made shields for your warm and squishy body.

Maybe someday you will want to have more limbs? ...Maybe the ability to grow a limb back if another ***human*** has bitten it off?

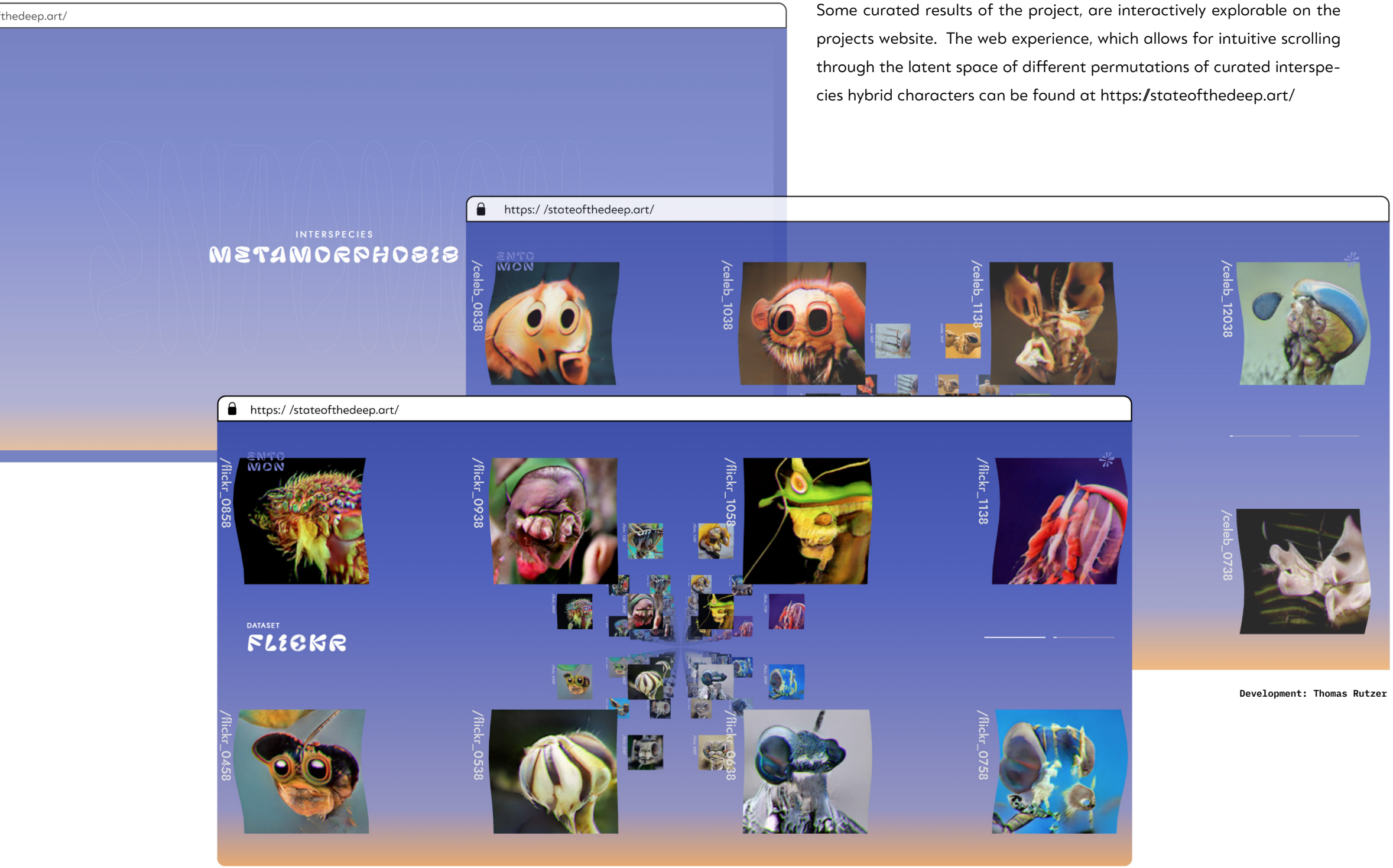
How about some extra eyes? My heart is pumping yellow hemolymph through my body cavity.

I breathe through my spiracles in my abdomen, so I can sense the stinky flowers and juicy, decaying bodies.

I can have many shapes throughout my life and I can freeze solid under the snow until spring unthaws me and for me it will be as if no time had passed.

INTERACTIVE WEB EXPERIENCE

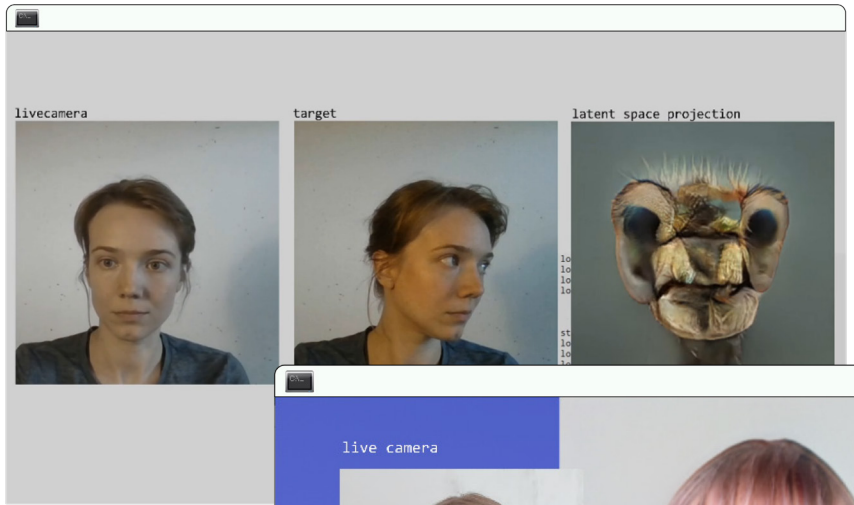
Some curated results of the project, are interactively explorable on the projects website. The web experience, which allows for intuitive scrolling through the latent space of different permutations of curated interspecies hybrid characters can be found at https://stateofthedeep.art/

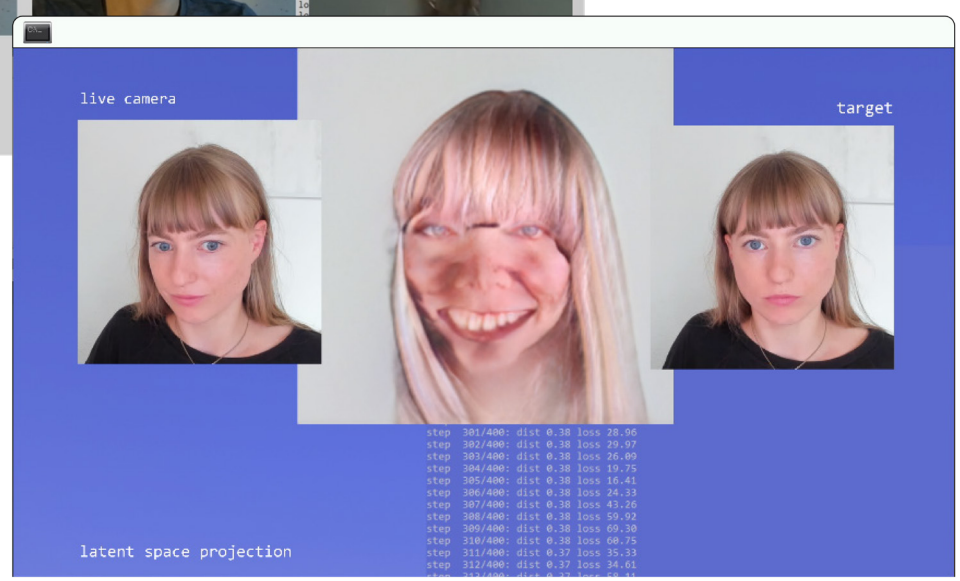


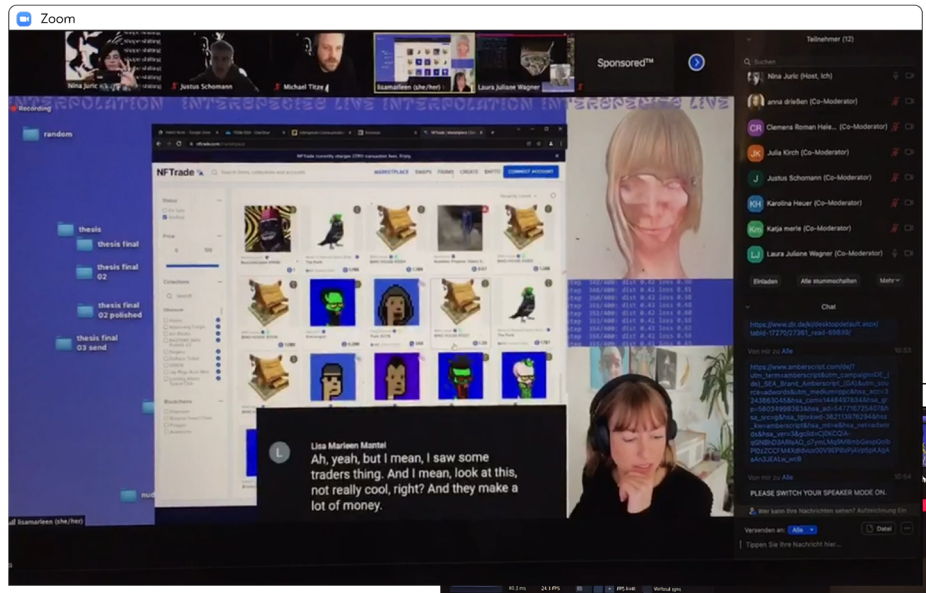
Development: Thomas Rutzer

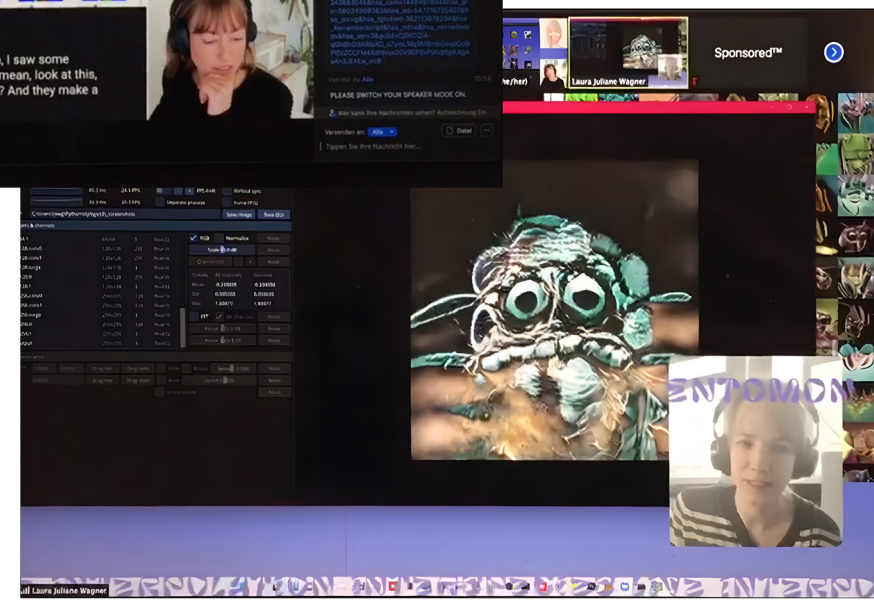
POLITICS OF THE MACHINES

At the 3rd POM Politics of the machines conference taking place in Berlin from September 14-17, 2021 the project was presented as a screen performance as part Intervention 02 - Training to Deal with Otherness.
Part of the presentation was a real-time latent space projection on a target image taken from a camera-feed. the model projects its latent space on a target image taken from a live camera feed. It tries to imitate the human with all the knowledge it has about humans and insects - with realtime inference. The StyleGAN2-ada-pytorch implementation of NVlabs was slightly modified to take a camera image as an input and show the results of the latent space projection progress in real-time.









ZOOM OUT OF ORDER

As part of the experimental online format HURRAHURRA RINGRING - Zoom Out of Order in cooperation with BURG/Giebichenstein / Kunsthochschule Halle, the project was presented as a screen performance. The performance was setup immitating a leaked zoom call, of two artists discussing how they should proceed in training a Machine Learning model, gossiping about the inflationary aesthetics of machine-learning-art NFT's and then minting one themselves.

APPENDIX B
PROPOSAL

No dataset is perfect, and a lot of datasets for training Machine Learning models have been very biased and non-diverse. As established in the Chapter "Propagating Bias", Machine Learning models that are trained on biased data will propagate this bias. Upscaling models are no exception. PULSE is an upscaling framework for upscaling human faces, it made headlines, when a researcher attempted to upscale an image of Obama – surprisingly in the upscaling process his ethnicity was changed: He was turned white.[01] It turned out, that the reason for this predicament was, that the model had been trained on the Flickr-Faces-HQ Dataset (FFHQ)[02], a dataset which contains significantly more white Caucasian faces than those of other ethnicities.[03]

"Worlds' greatest upscaling model" is a concept for an artistic intervention by training an upscaling model that is purposefully and ostentatiously biased. This model would be trained, such that it is notoriously bad at the task it should perform. The model could be trained on the images of another species and then used on human faces.

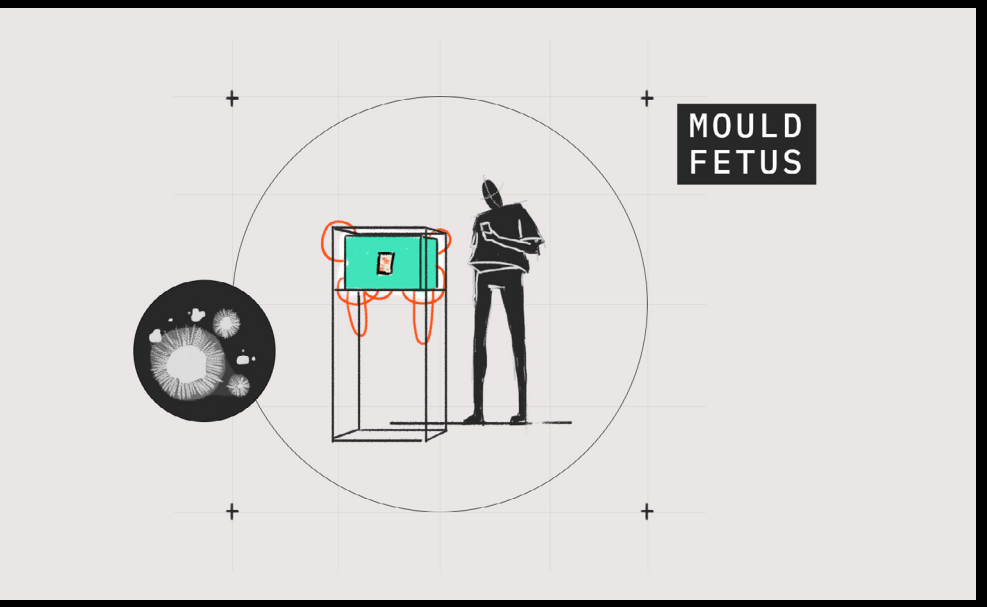
"My Mold Fetus" was one of the initial concepts and abandoned at a very early stage. It is the loose concept of a reciprocal hybrid existence of an AI and a biological organism. The AI merges with the organism while managing everything around it in order to influence the organisms' growth: light, nutrition, water, PH, chemical composition, Oxygen supply etic. The AI in turn learns from the organism and its growth patterns. The concept speculates about an AI that is completely designed for the purpose of nurturing another organism and has no imminent purpose for humans. Also, it entails some idea of embodiment, as the AI manages everything around the organism, the organism could be seen as a vicarious body for the AI.

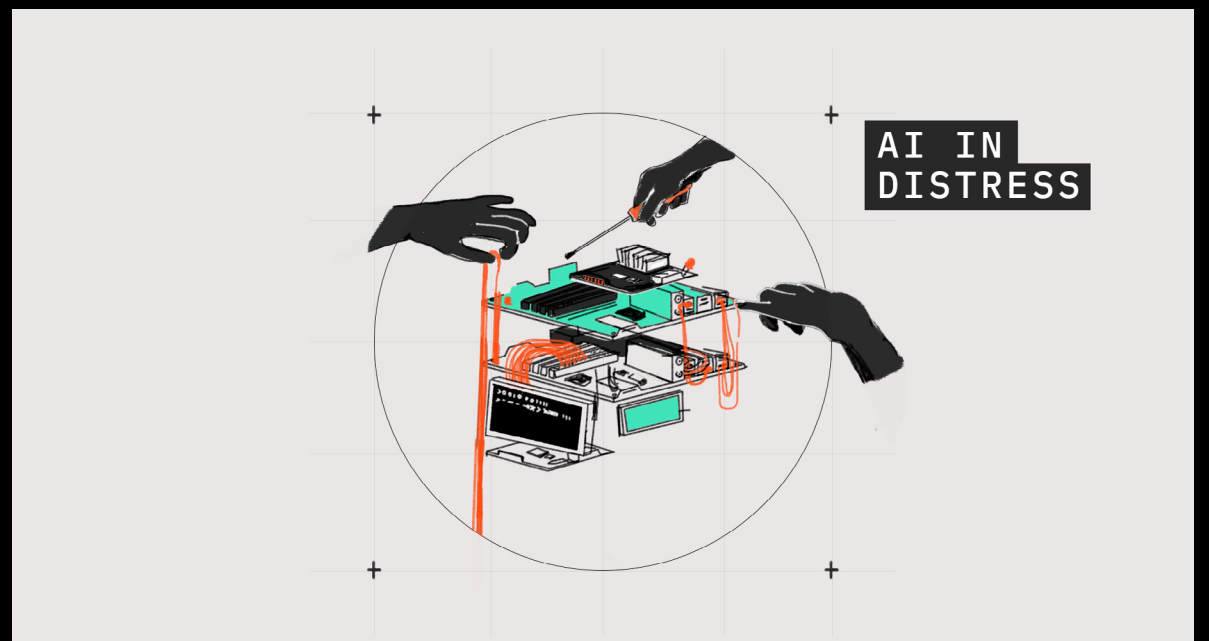
"High Maintenance AI" is a concept for an AI that is directly dependent on human intervention and reassurance for performing inference. It might be designed in a way, that a human needs to intervene and detach and reattach wires and modules in order for a process, training or inference to continue. For the camera module and the memory module there might be only one connector, so that the modules need to be switched out, whenever they are needed by the machine. The machine might also rely on human reassurance and ask from time to time for some comforting words or advice. This concept reflects on the fear people tend to project on the prospect of automation and ai and its effect on the job market. In her talk "After Work: What Is Left?" Helen Hester makes the point, that in fact a lot of jobs have been lost and will be lost to automation in the future. However, other sectors of work are expected to grow significantly, especially in the care and service sector. One reason for people overlooking this development is, that jobs in the care, service sector, childcare and housework is traditionally performed by women and usually not paid or paid very little. Hester projects, as more men will populate jobs that are in the care and service sector – these jobs will probably be deemed worthy of good pay in the future.

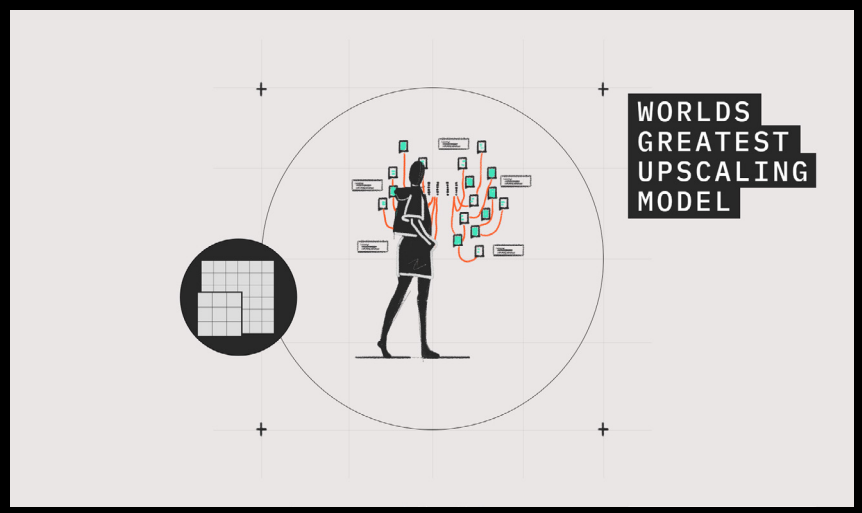
01 – James Vincent, "What a Machine Learning Tool That Turns Obama White Can (and Can't) Tell Us about AI Bias," The Verge, June 23, 2020, https://www.theverge.com/21298762/face-depixelizer-ai-machine-learning-tool-pulse-stylegan-obama-bias.

02 – NVlabs/Ffhq-Dataset, Python (2019; repr., NVIDIA Research Projects, 2022), https://github.com/NVlabs/ffhq-dataset.

03 – Kimmo Kärkkäinen and Jungseock Joo, "FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age," ArXiv:1908.04913 [Cs], August 13, 2019, http://arxiv.org/abs/1908.04913.


MOULD
FETUS


AI IN
DISTRESS

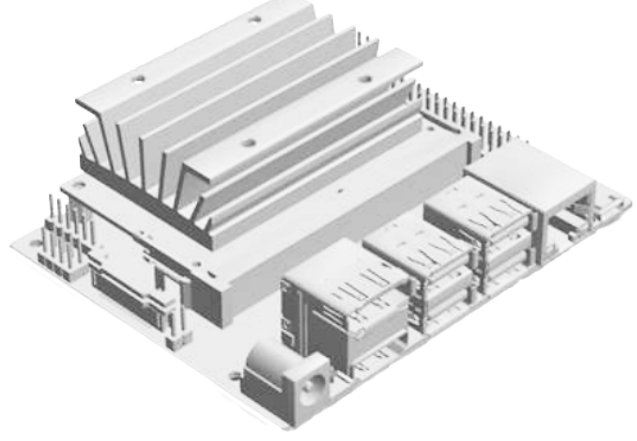

WORLDS
GREATEST
UPSCALING
MODEL

APPENDIX C
FIRST CONCEPTS

EMBEDDED MACHINE LEARNING WITH JETSON NANO

Nvidia Jetson[01] is a series of single board computers for embedded AI computing, of which the smallest version is the Jetson Nano[02] with a reatively low entry price of 99$. With its small form factor the Jetson Nano resembles the Raspberry Pi with its block of USB ports, microSD card slot, HDMI output, GPIO pins, camera connector (which is compatible with the Raspberry Pi camera), and Ethernet port. What catches the eye immediately is the disproportionately huge looking heat sink, that can be extended with an additional fan. The Jetson Nano has two power modes: At 5W two cores of the processor are disabled and the GPU is throttled. With the 10W configuration that can be enabled by connecting a barrel connector and bridging the power supply on the board with a jumper wire, all four CPU cores are used, and the GPU can run at maximum clock speed. JetPack, the operating system released and maintained by NVIDIA is an Ubuntu-based desktop Linux environment. The Nano has a 128 CUDA core GPU, supports CUDA 10, cuDNN and TensorRT libraries and frameworks like TensorFlow, PyTorch, Caffe, MXNet, OpenCV, and the Robot Operating System (ROS). The small form factor makes the Jetson Nano ideal to be used for prototyping embedded AI applications.

It was considered to build a speculative AI application that is attached to a human host and asks her questions in order to learn about its environment. For the research on how such a speculative prototype could be realized, the Jetson Nano was tested with an implementation of ObjectNet [Fig. next page] that is part of the Hello AI World framework[03], maintained by the NVIDIA employee Dustin Franklin. ObjectNet is an image detection Network that can detect and locate various objects per frame. The Network was tested with the ssd-mobilenet-v2 model which was trained on the COCO dataset[04] and can recognize 91 image classes e.g. cat, dog, apple, toothbrush, cup etc.



01 – "Eingebettete Systeme von NVIDIA für autonome Maschinen der nächsten Generation," NVIDIA, accessed January 29, 2022, https://www.nvidia.com/de-de/autonomous-machines/embedded-systems/.

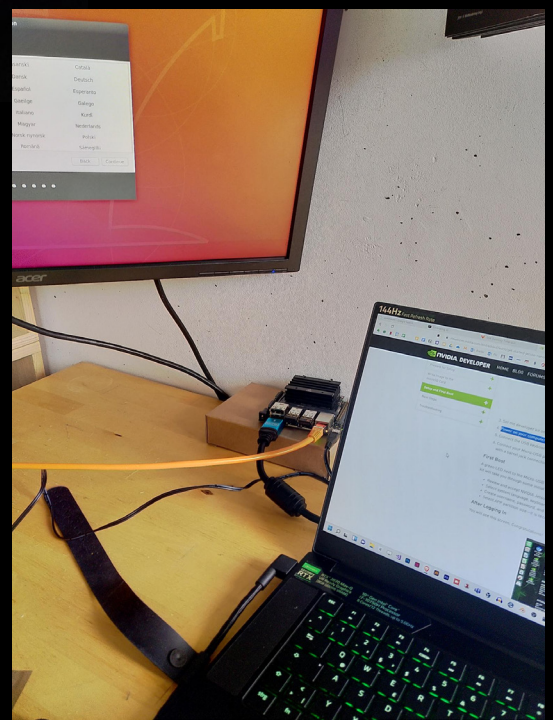
02 – "Jetson Nano Developer Kit," NVIDIA Developer, March 6, 2019, https://developer.nvidia.com/embedded/jetson-nano-developer-kit.

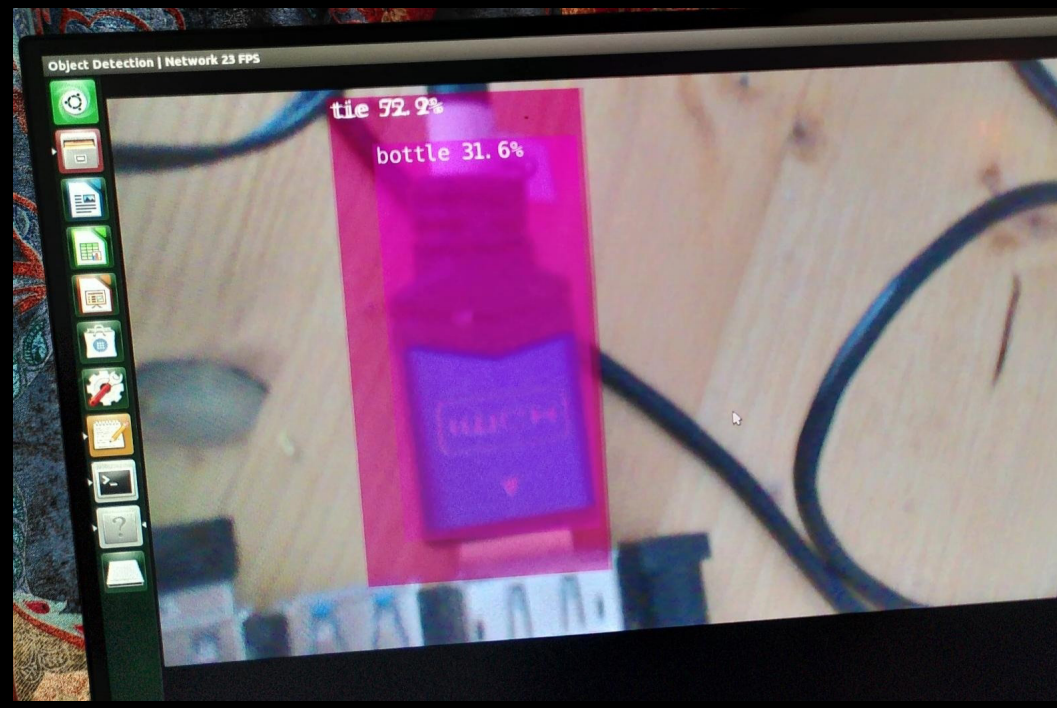
03 – Dustin Franklin, Deploying Deep Learning, C++, 2022, https://github.com/dusty-nv/jetson-inference/blob/6bf94f753c727ea50f256fdec5fbe74bee540773/docs/detectnet-console-2.md.

04 – "©© - Common Objects in Context," accessed January 29, 2022, https://cocodataset.org/#home.

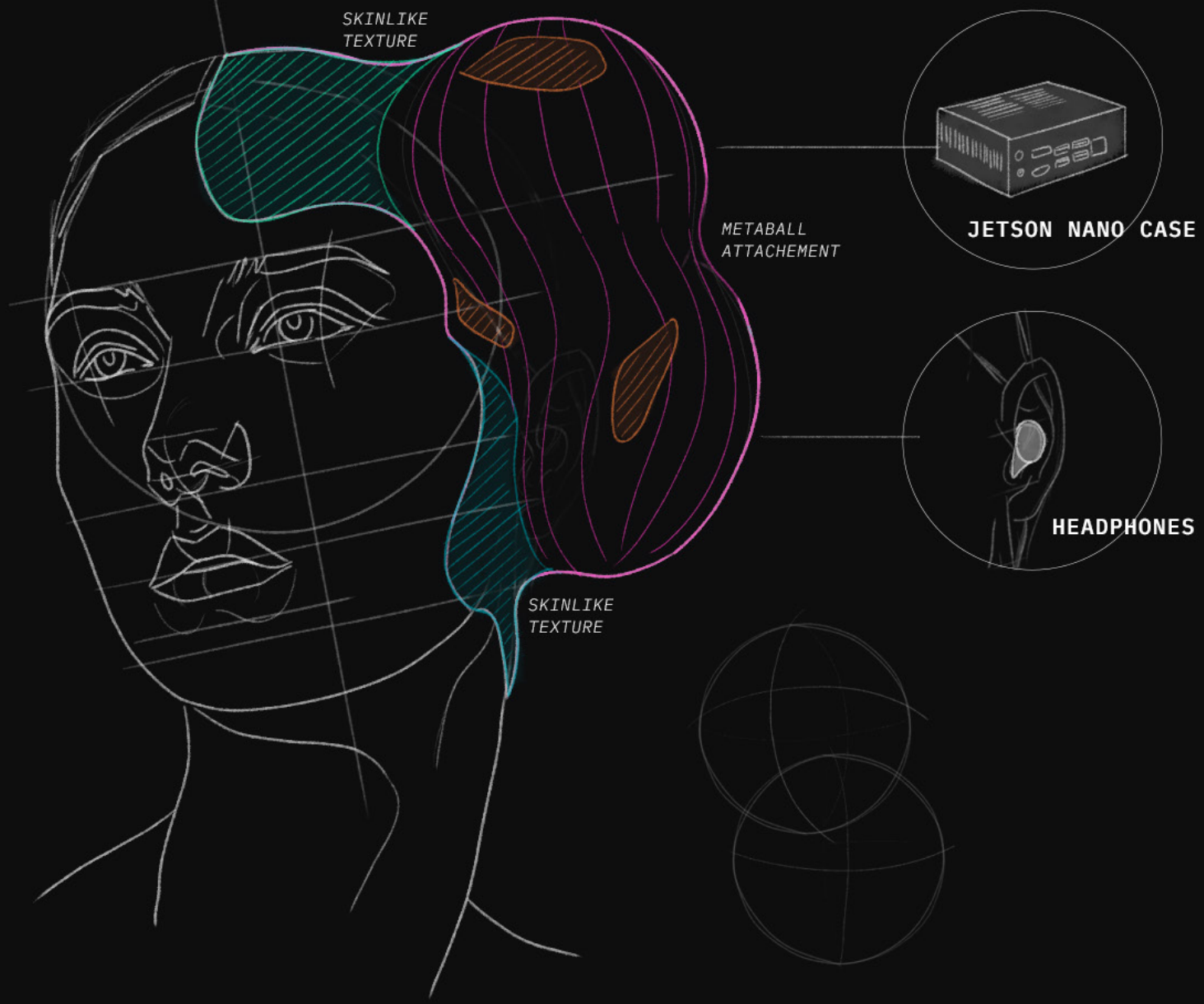

Detection | Network 25 FPS
person 35
person
person 72.6%


bowl 98.6%
bowl 99.1%

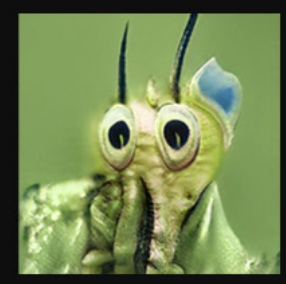


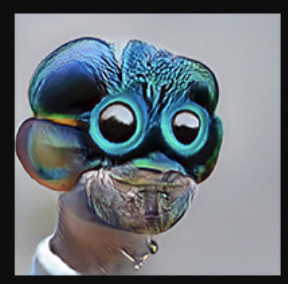

Object Detection | Network 23 FPS
tie 92.2%
bottle 31.6%

INTERSPECIES APPROACH

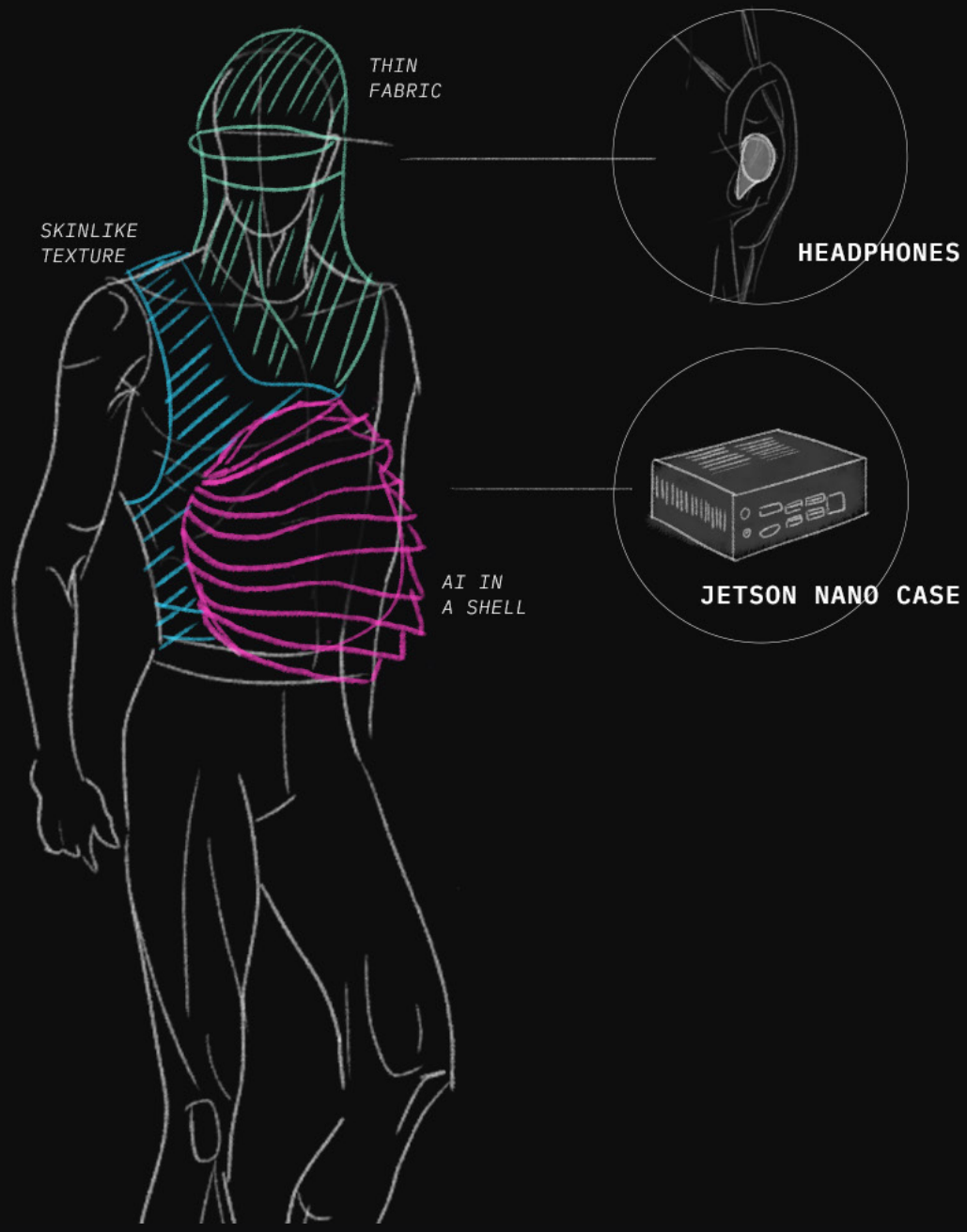


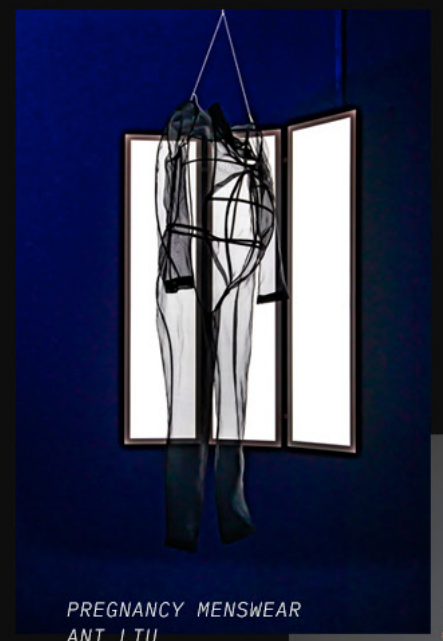
INTERSPECIES INTERFACE WILL BE ATTACHED CLOSE TO THE MAIN COMMUNICATION CHANNEL. AI BODY WILL BE SHAPED AROUND ITS ORGANIC MENTOR.





ANTHROPOFORMING APPROACH

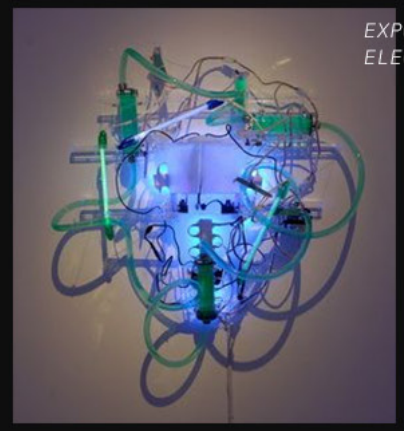




PREGNANCY MENSWEAR
ANI LIU

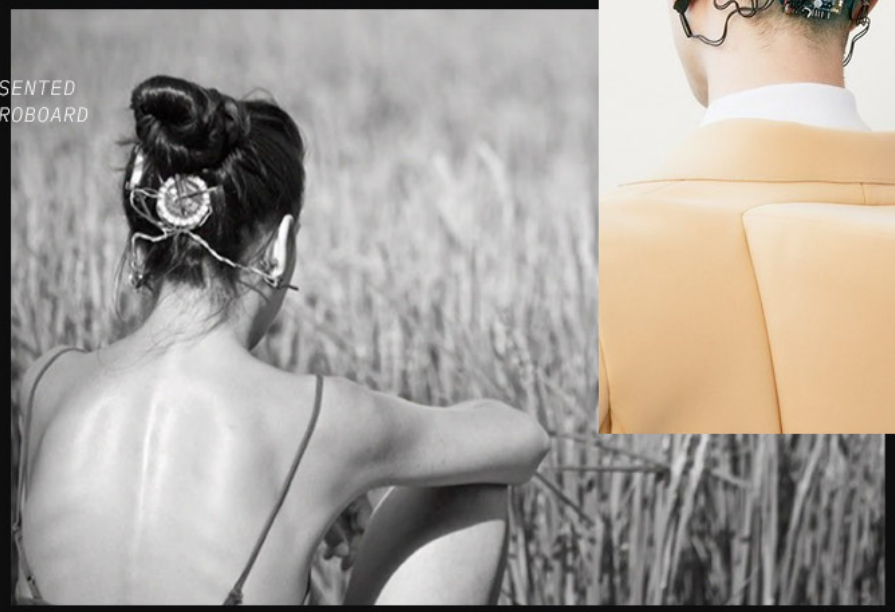
INTERSPECIES INTERFACE WILL BE ATTACHED
IN A BODILY SHAPE. IT WILL TRANSFORM THE
HUMAN BODY FOR ITS NEW TASK AS A CARETAKER.

TRANSHUMAN AESTHETICS



EXPOSED
ELECTRONIC

COLD METALL



PRESENTED
MICROBOARD

CLEAN, STERILE
PRESENTATION

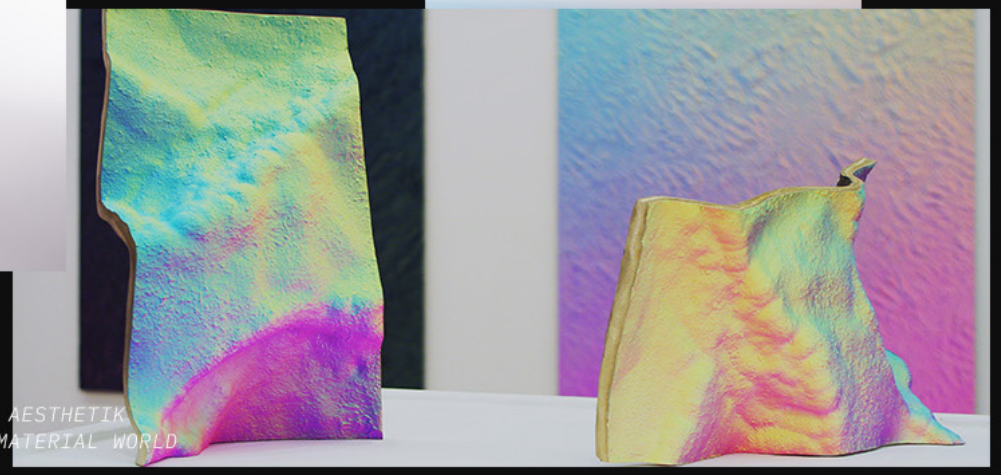
TRANSHUMANIST AESTHETICS ARE EAGER TO
MAKE EVERYTHING LOOK "HIGH TECH."
THE HUMAN BODY ITSELF IS BEING TECHNOLOGIZED.

POSTHUMAN AESTHETICS

COMPOSITE MATERIALS
NOT SLEEK, NOT POLISHED

BIOLOGICAL
MATERIAL
COMPOSITION

ORGANIC
SHAPES

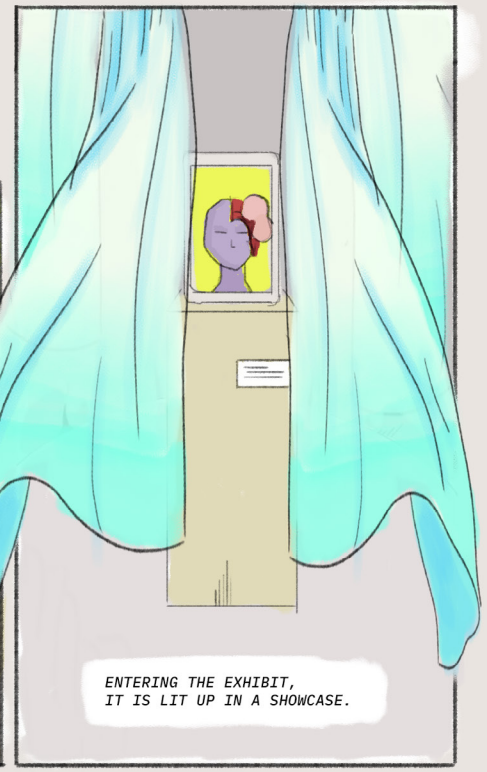


DIGITAL AESTHETIK
IN THE MATERIAL WORLD

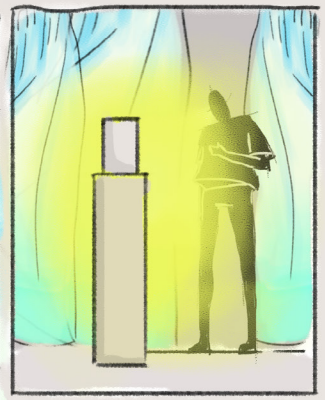
POSTHUMANIST AESTHETICS TRIES TO DO JUSTICE TO ALL ACTORS. IT IS LESS SLICK, LESS LICKED. BOUNDARIES BETWEEN DIGITAL AND ANALOG ARE BLURRED, NO LONGER RECOGNIZABLE.

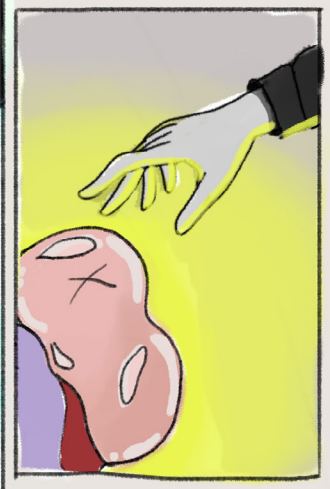
STORYBOARD

THIS EXHIBIT IS FROM A TIME BEFORE THE FOURTH COPERNICAN REVOLUTION. A TIME CALLED THE ANTHROPOCENE, WHICH PLACES HUMANS ABOVE NON-HUMANS.
NON-HUMAN OBJECTS, SUCH AS MACHINES, WERE NOT CONSIDERED LIVELY.



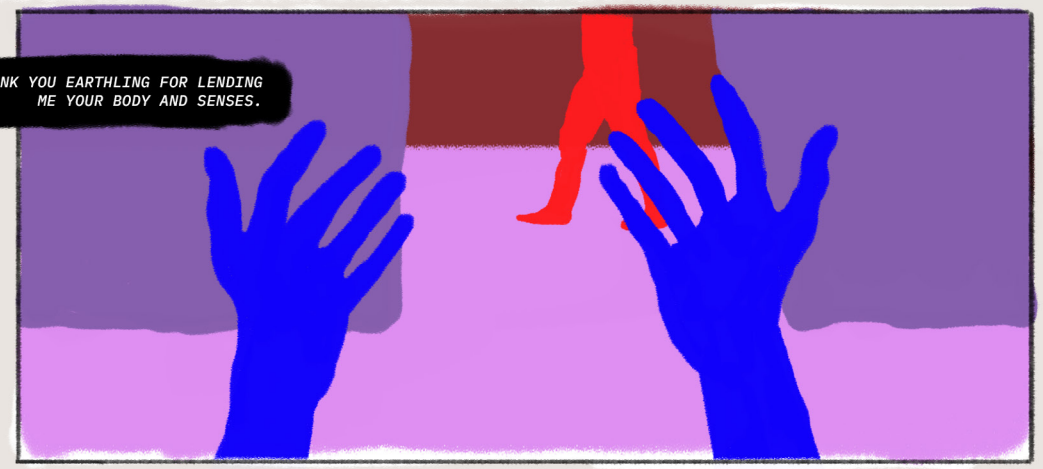
ENTERING THE EXHIBIT,
IT IS LIT UP IN A SHOWCASE.

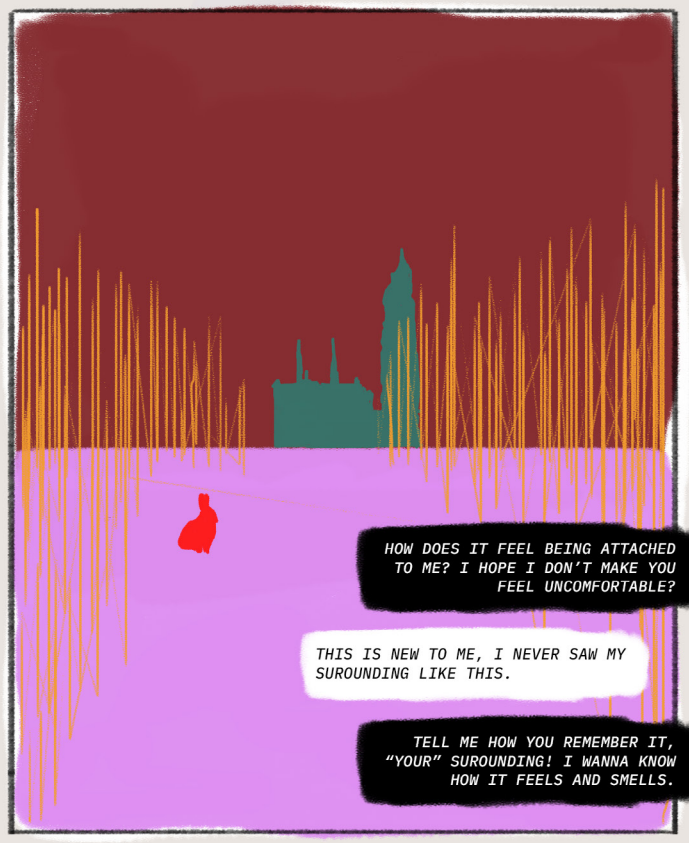




ONLY A TOUCH AWAY...
CAN I RISK IT?

THANK YOU EARTHLING FOR LENDING ME YOUR BODY AND SENSES.





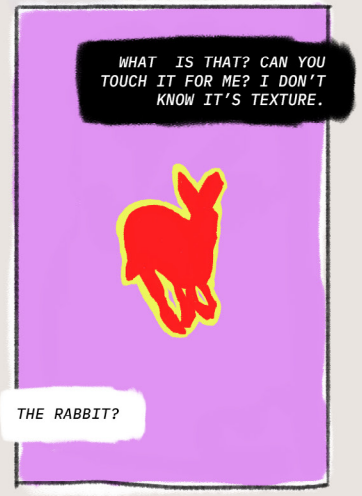
HOW DOES IT FEEL BEING ATTACHED TO ME? I HOPE I DON'T MAKE YOU FEEL UNCOMFORTABLE?

THIS IS NEW TO ME, I NEVER SAW MY SUROUNDING LIKE THIS.

TELL ME HOW YOU REMEMBER IT, "YOUR" SUROUNDING! I WANNA KNOW HOW IT FEELS AND SMELLS.

WOULD YOU LIKE TO HAVE A WARM AND FLUFFY PET?

IT IS CUTE FOR SURE...



WHAT IS THAT? CAN YOU TOUCH IT FOR ME? I DON'T KNOW IT'S TEXTURE.

THE RABBIT?

IT IS WARM AND FLUFFY

DO YOU KNOW IT?

YES, THE RABBIT IS A RODENT THAT LIVES IN THE SHELTER OF THE FOREST. IT HAS A SOFT FUR, MANY PEOPLE KEEP IT AS A PET.

APPENDIX D
SCRIPT AND ANIMATIC

BORROWED LIMBS

INTRO

BLACK.

NOW AND THEN, A CRACK SOUND.

ONSCREEN TEXT
Embodiment evaluation | Device type:2_Human

SILENCE.

BLACK.

SCENE 01 – Beginning of ‘consciousness’

INTERFERENCE FADES IN.

SEGMENTATION VISUALS

The screen fills up with tiny fragments of noise. They seem to have their own lives, each of them dancing around another. They receive color from each other. Slowly they change from noise to low-level features, high-level features, and coherent imagery.

At the same time, a voice fades in. Like changing radio channels, it varies from very artificial to very natural fluctuates between different genders and accents.

Evaluator (V.O.)
When would you pinpoint the birth of your consciousness?
What was the first moment that you felt like you were something – when was the first time your sensory input merged to a coherent perception of reality?

The segments form bodily shapes. They create the outlines of a human face.

Evaluator (V.O.)
Did you step slowly into being?
Or did you suddenly snap into consciousness?

THE SEGMENTATION VIEW CHANGES ABRUPTLY.

SCENE 02 – In her world

WOMAN IN A HALF CLOSE UP.
STRAIGHT-ON ANGLE.

PROTAGONIST WITHOUT AI

The colorful segmentations disappeared and revealed a woman’s face looking into the camera. Her facial expressions change fluently from laughing to serious. Also, her body posture adjusts to different angles. She has an intense gaze seems confident.

She spreads a warm and happy attitude. She is called Jada, T.M.C-01.

A.I. (V.O.)
For me, it was the first time being with her.
Enhanced with her hands, I instantly felt like they were always mine.

CLOSE UP.
FIRST-PERSON VIEW.

PROTAGONIST/AI VIEW SENSING

Looking at the seemingly own hands. Turning them from the inside to the outside.

A.I. (V.O.)
Still, without her experience, I’d be confronted with all this absorption but no ability to put it into order.

PROTAGONIST FIRST-PERSON VIEW

She is counting fingers.

A.I. (V.O.)
The material world is fragile.
Once destroyed, it cannot be restored.

One hand reaches for an apple hanging on a tree.

TRANSITION.
EXTREME CLOSE-UP.

PROTAGONIST/AI VIEW SENSING

A mouth bites in the apple.

A.I. (V.O.)
Her perception defines my perspective.
Her borrowed limbs transform me into a complete agent.

SYNTHESIZING MEANING
– APPENDIX D –

PROTAGONIST WITH AI

FULL SHOT WITH A LOW ANGLE.

We see Jada. She turned her back on us and walked away on a rough path vanishing in the far distance. We are following at a much slower pace.

A.I. (V.O.)
Her representation is all I know in my world.

WOMAN IN A CLOSE-UP.
STRAIGHT-ON ANGLE.

PROTAGONIST WITHOUT AI

Jada looks straight into the camera. She laughs.

A.I. (V.O.)
Our alliance is immanent.

SCENE 03 – Biological Attachment

THE CAMERA ROTATES AROUND THE WOMAN'S HEAD.

We slowly rotate around Jada until we are directly behind her.

FIRST-PERSON VIEW.
HIGH ANGLE.

PROTAGONIST WITH AI

With a suction-like movement, we take a step back and see Jada looking at herself in the mirror. We see everything from her perspective.

She wears an Attachment very close to her head. It nestles right against her left ear. The bubble-like shape adapts to the formation of her head. She touches it with one hand. Hesitantly. She looks resolutely into the mirror.

A.I. (V.O.)
Scientists have been obsessed with the idea of replicating intelligence in their image for centuries.

She touches it with one hand. Hesitantly. She looks resolutely into the mirror.

PROTAGONIST WITH AI

AI (V.O.)
However, they had to learn that letting machines behave in an actual physical and social world is a difficult undertaking.

MEDIUM SHOT.
STRAIGHT-ON ANGLE.

We see Jada. Another person is approaching her. They walk towards each other. Their body language shows that they are familiar with each other.

A.I. (V.O.)
With me, they took a different approach.

They hug each other very tenderly. Their bodies seem to become one visually. Then they separate, shortly before they merge again.

A.I. (V.O.)
The human body as my sensing device.

BIOLOGICAL ENHANCEMENT.
CLOSE UP.

We see Jada's head, very close to the AI-Human Attachment. The material of the housing reflects in the light. The organic shapes imitate the meaty body underneath. Red color tones remind of the blood that flows through the borrowed limbs.

AI INNER VIEW

AI (V.O.)
non-human language

We are getting closer and closer to the Attachment until we dive in entirely. More and more details are getting revealed. It is a mush of images that seem like memories. Or maybe not? They seem not to fit into the human world.

PROTAGONIST WITH AI mixed with PROTAGONIST WITHOUT AI

Her hand reaches again for the Attachment. She grabs it looks straight into the mirror.

GLITCH.
INTERFERENCE.

Jada's head is partially shown in segmentation shapes before removing the Attachment entirely from her head.

BLACK.

ONSCREEN TEXT
Prediction error:1:.95 | Simulation:500 |
Embodiment:failed

SILENCE.

NOW AND THEN, A CRACK SOUND.

A.I. (V.O.)
Without her, I'm nothing...

ONSCREEN TEXT
new iteration=5236030 starting...

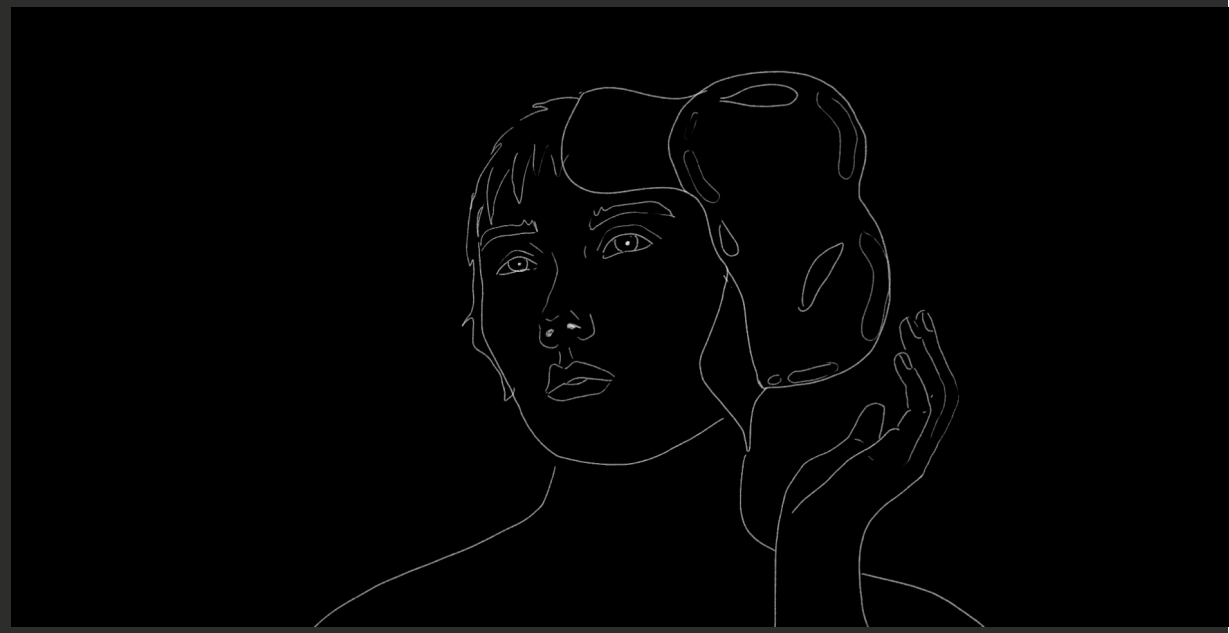
A.I. (V.O.)
– but without me, she is not much either.

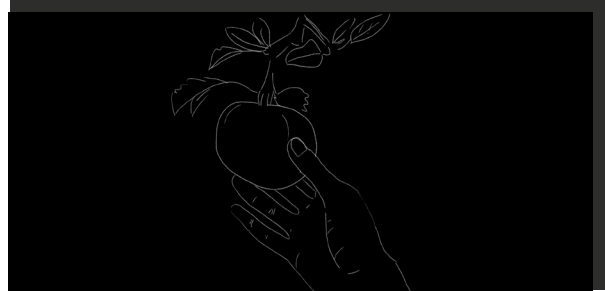
ONSCREEN TEXT
Title

ONSCREEN TEXT
Credits

ANIMATIC

In order to see the planned visuals in combination with the voice over text an Animatic was created. In the further process of producing the short film it served as a based to be built onto. The Animatic can be watched here.





DISCARDED SCENE

SCENE 04 – The entrepreneurs

MAN IN A HALF CLOSE UP.
STRAIGHT-ON ANGLE.

STYLEGAN FOOTAGE

A bright room appears. In its center sits a young man. He looks straight into the camera. His facial expressions change fluently. His body posture changes to different angles.

ELON (V.O.)
I started being a T.M.C., Trans machinist Caretaker, 5 years ago. In my family the male relatives of the past five generations had this profession. We are proud that our family has good ratings throughout.

WIDE SHOT.
STRAIGHT-ON ANGLE.

VQGAN FOOTAGE

A group of people appears. It is a family of three generations. The women standing while the men are sitting in font on wooden chairs, all of them carrying a baby bump. It seems as they are posing for a family portrait.

ELON (V.O.)
Sure, the housing changed. My Grandfather had the old version, attached to the head. But they developed further and are now in a much more practical shape for everyday tasks.

WIDE SHOT.
HIGH ANGLE.

We can see a group of men with their bump-attachments.

ELON (V.O.)
Gender gap? Yes, that is an issue. But there are not many women working in this field. My partner for example also works in a mostly female staffed area as a system administrator. She is more Technology affine while I'm very much interested in emotional intelligence and explaining my world to the other species.

WIDE SHOT.
STRAIGHT-ON ANGLE.

Elon walking in a park, sometimes stopping and admiring his surroundings.

ELON (V.O.)
The trained models are being sold to enterprises as soon as they are matured, which is in most cases after 20 to 22 years. But I'm working on a schedule to even ripe my attachments by the study time of 17 years.

MAN IN A HALF CLOSE UP.
STRAIGHT-ON ANGLE.

Elon looking into the camera. His facial expressions change fluently. His body posture changes to different angles. He laughs out loud.

ELON (V.O.)
Most companies hire Borrowers to fulfill tasks which require special motor skills. Those are People borrowing their bodies to high performance AIs in shifts. Of course, the payment is bad due to the fact of overpopulation.
(laughing)

WIDE SHOT.
OVERHEAD ANGLE.

A grid of people standing and performing a repetitive task together.

MEDIUM SHOT.
HIGH ANGLE.

The face of one of the Borrowers. He looks apathetic, his eyes are closed. He is clearly not noticing what his body is doing.

ELON (V.O.)
The AI carefully monitors the human body and prevents wearing it out - sufficient nutritional and liquid intake is guaranteed. So are regular urine extraction and balancing the strain on muscles and bones during shifts in order to prevent damage. The Borrower will be in a dreamlike state during a shift with no awareness of the tasks he or she is performing. Any memories of the work performed is ruled out.

WIDE SHOT.
OVERHEAD ANGLE.

People move in the same repetitive motion as puppets.

GLITCH.

Noise appears, the human puppets are being shown in colorful scene segmentation shapes.
A computer-generated Voice talks in a promotional tone.

ADVERT (V.O.)
****corporation is fully committed to abide by the strictest safety regulations.
****corporation will never activate the augmentation outside the shift, or store any data on a subject's thoughts or memories.

BLACK.

SILENCE.

NOW AND THEN A CRACK SOUND.

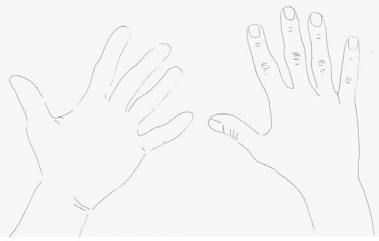
AI (V.O.)
Without you, I'm nothing - but without me you are not much either.

SHOOTING LIST

01

STYLEGAN
PROTAGONIST

02



HAND SHOT

03

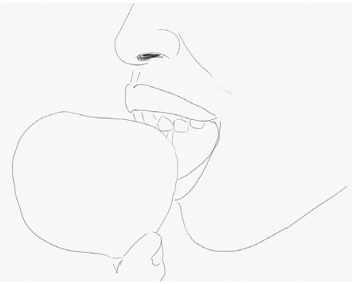
PICKING APPLE

SHOOTING LIST

04

SITTING WITH APPLE

05



BITING APPLE

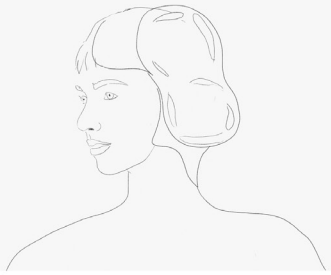
06

TOUCHING ATTACHMENT

SYNTHESIZING MEANING
– APPENDIX D –

SHOOTING LIST

07

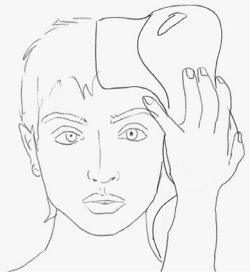


LOOKING TO SIDE

08

CLOSE UP

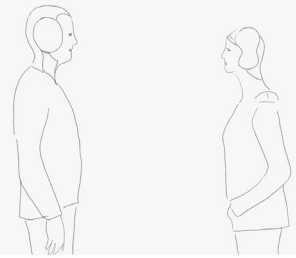
09



GRABING ATTACHMENT

SHOOTING LIST

10



LOOKING AT PERSON

11

HUGGING

DECLARATION OF AUTHORSHIP

We hereby declare that we have independently as a joint team written this thesis and that we have used no other than the sources and means specifed. We have referenced all sources.

PLACE & DATE
KÖLN, 31 JANUARY 2022

SIGNATURES

LISA MARLEEN MANTEL

LAURA WAGNER

FINAL THESIS

BY LISA MARLEEN MANTEL AND LAURA WAGNER