# Documentation

"Behind all of the manifestations of the eerie, the central enigma at its core is the problem of agency. In the case of the failure of absence, the question concerns the existence of agency as such. Is there a deliberate agent here at all? Are we being watched by an entity that has not yet revealed itself? In the case of the failure of presence, the question concerns the particular nature of the agent at work."[[1]](#footnote-1)

This text intends to give insight into everything that was and is happening behind the scenes of this project. On the technological side, this project consists of multiple components, most of which are based on so-called "artificial intelligence" (short AI). All messages and selfies are being generated live. The first half of this text will give a (critical) overview of this technology. The second half will go into the process of creating this project and explain / go into detail about my choices and motivations.

## Machine Learning

Describing machine learning[[2]](#footnote-2) is challenging. This is not necessarily because it is a complicated technology. Most technologies were perceived to be complicated when they were new and yet were rather easy to explain once humans had gotten used to them and had a common vocabulary to describe them. While AI’s use outside of research is still a recent development, most people have now had encounters with AI, knowingly or unknowingly. A common vocabulary seems to be developing.

Exactly this emerging vocabulary makes describing machine learning challenging, though. A common description of it might be: "Machine learning describes algorithms that are able to 'learn' certain patterns through being 'shown' vast amounts of data. These patterns could be anything: visual patterns in images, grammatical patterns in text, etc. If you want an AI to be able to 'discern' sad and happy facial expressions, for example, you could show the algorithm tens of thousands of pictures of sad and happy faces that are labeled as such. After a while of training with these pictures, the AI will be able to put the labels 'sad' and 'happy' on completely new portraits that it has never ‘seen’ before. Aside from labeling or classifying data, machine learning algorithms can also generate new data. An algorithm that is 'trained' with faces, for example, will 'learn' what a face looks like and will be able to construct completely new faces with that knowledge."

The problem with this explanation is that, while machine learning algorithms can label or generate data through a process called training, they are probably not capable of "learning" or "discerning" anything in any way that would be familiar to humans, although these terms might suggest otherwise. As Matteo Pasquinelli and Vladan Joler put it: "Machine learning is a term that, as much as 'AI', anthropomorphizes a piece of technology: machine learning learns nothing in the proper sense of the word, as a human does."[[3]](#footnote-3)

It is important to keep in mind that, as much as the field of AI is a technological praxis, it is also a narrative one. In the words of Phil Agre: "As a practical matter, the purpose of AI is to build computer systems whose operation can be narrated using intentional vocabulary. Innovations frequently involve techniques that bring new vocabulary into the field: reasoning, planning, learning, choosing, strategizing, and so on. Whether the resulting systems are really exhibiting these qualities is hard to say."[[4]](#footnote-4)

This blurring of concepts between humans and machines may be in part responsible for the mystification that is happening around AI. Even the expression “artificial intelligence” implies some kind of autonomous, mystical, alien computer minds.[[5]](#footnote-5) The idea of thinking (and maybe even feeling) machines is, of course, an old one to be found in literature and films. So naturally, in this time of transition where AI is seeing more and wider use, there are still sci-fi-like stories floating around in our minds. The question of agency is at the core of many of these stories where often an AI becomes self-aware and sometimes devious. This may be part of the eeriness that is still present in our perception of AI.

According to Mark Fisher "the sensation of the eerie occurs either when there is something present when there should be nothing, or is there nothing present when there should be something".[[6]](#footnote-6) For most of us, there is something eerie in the fact that machines seem to display cognition or even opinions and feelings (in the case of text-generating algorithms) when they are just machines and as those, should not be able to think or feel anything. After working for a while with AI this impression turns upside down. Now it feels eerie that intelligence is perceived to be there when there is actually nothing like it present.

## Choices

In the end, the questions of whether machines can think were not of interest to me in this project, but the matter of what humans project into machines was. The eeriness that is perceived reminded me of ghost stories, especially when thinking about the question of agency. Is something "present"? Can it "see" me? The proper answer, laid out in the text above, is: "It depends on what expressions like 'present' and 'see' mean to you, but probably not." This is interestingly a similar answer that I would give to somebody, who asks me, whether ghosts exist. Yet I certainly know the feeling of having a shiver run down my spine because I am in an eerie place and I think of its history.

I am not the first to make the connection between AI and ghosts, however. Business got there way before me. As an example: In 2014, the start-up eterni.me was founded with the aim of enabling people to exist beyond their death: By handing over their most intimate data as well as access to their social media accounts, AIs are supposed to be able to create avatars after the death of these people, which "live on" ghostly in their place.[[7]](#footnote-7)

This leads into another ghostly quality to AI: The fact that it can never truly generate anything new. As it is able to regenerate and predict patterns from its training datasets, it is bound to the information in that data. Thus "machine learning automates the dictatorship of the past, of past taxonomies and behavioral patterns, over the present. This problem can be termed the regeneration of the old […]".[[8]](#footnote-8) One could say, rather dramatically, that machine learning algorithms are doomed to echo past events, voices, and faces (anything that it's trained on, anyway). Rather like one might imagine a ghost: Being an echo of someone who once lived and is now neither really living nor fully dead, but constrained to reenact what its living self once did and said.

All of the above motivated me to create this “spooky” interface, where text and messages shift around and which always seems a bit too slow to be an actual working chat app. With the messages and selfies, too, I meant to work with the eerie, skating around the questions of presence and absence. Sometimes they make total sense: A conversation unfolds, that could be between humans and the images look almost as if they'd actually been taken by a camera. Then, suddenly, all this falls apart again, when grammar fails or a part of the conversation gets repeated over and over like a broken record.

## Process

Every machine learning model needs a dataset. In this case building the dataset started with e-mail conversations I had with Hendrik Kempt, a philosopher, working in applied ethics. These conversations were topic driven, revolving around the use of AI, but also digital surveillance and control, ghosts and the way we think as scientists, artists and programmers. Later I rewrote these e-mails into a chat and added an improvised conversation between two performers. I used the latter to guide the conversation further and add emotional content, but also to anonymize the results somewhat.

Here I was using the ability of machine learning to compress the dataset during training.[[9]](#footnote-9) As the conversation between Hendrik and me and that between the performers gets put into the same dataset, it will generate a mix of both conversations, sometimes leaning more into one style and topic of conversation and sometimes more into the other. I did the same with the selfies, conflating my face and one performer’s together and Hendrik’s face and the other’s.

For the sounds that get played, when a message gets received, I worked differently: I recorded all default notification sounds of my phone and used an algorithm to compress them into one sound. The result is a list of numbers. Now I put slight variations on those numbers and used the same algorithm to turn them back into audio files. The result is that each sound is now abstracted randomly.

I need to add that I did not train any machine learning model from scratch. For this, I would have needed several libraries worth of text for the messages and at least 50.000 images for the selfies. Instead, I used a technique called "finetuning". For finetuning one uses models that were already trained[[10]](#footnote-10) and trains them for a little longer with a much smaller, but more focused, dataset. So the models were already able to create images and text respectively before I finetuned them, but through finetuning they now generate texts, that somewhat approximate my datasets. While it is unlikely that faces or words pop up that are completely different from my dataset, it is worth mentioning that the pre-training datasets are still compressed in there somewhere, maybe peeking out at opportune moments, to reveal themselves.

1. Mark Fisher. *The Weird and the Eerie*. Repeater 2017. P. 63. [↑](#footnote-ref-1)
2. The terms machine learning and artificial intelligence are often used interchangeably. Machine learning actually is a technology that is a subset of the field of AI. [↑](#footnote-ref-2)
3. Matteo Pasquinelli and Vladan Joler. *The Nooscope Manifested: Artificial Intelligence as Instrument of Knowledge Extractivism*. 12 June 2022. https://nooscope.ai [↑](#footnote-ref-3)
4. Phil Agre, *Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI*. 12 June 2022. https://pages.gseis.ucla.edu/faculty/agre/critical.html [↑](#footnote-ref-4)
5. Cf. Matteo Pasquinelli and Vladan Joler. *The Nooscope Manifested: Artificial Intelligence as Instrument of Knowledge Extractivism*. 12 June 2022. https://nooscope.ai [↑](#footnote-ref-5)
6. Mark Fisher. *The Weird and the Eerie*. Repeater 2017. P. 61. [↑](#footnote-ref-6)
7. Etnernime: Eternime Alpha Two, https://vimeo.com/142260863, abgerufen am 26.2.2021. [↑](#footnote-ref-7)
8. Matteo Pasquinelli and Vladan Joler. *The Nooscope Manifested: Artificial Intelligence as Instrument of Knowledge Extractivism*. 12 June 2022. https://nooscope.ai [↑](#footnote-ref-8)
9. Cf. Matteo Pasquinelli and Vladan Joler. *The Nooscope Manifested: Artificial Intelligence as Instrument of Knowledge Extractivism*. 12 June 2022. https://nooscope.ai [↑](#footnote-ref-9)
10. In the case of the messages, the model was pre-trained with the [PILE](https://github.com/EleutherAI/the-pile) dataset, and in the case of the selfies with [FFHQ](https://github.com/NVlabs/ffhq-dataset). [↑](#footnote-ref-10)