Generative AI & Synthetic Media

Rachel Baerts

Main question: What is generative AI?
Sub questions:
- What is the history of generative AI?
- How does generative AI work?
- Which different generative AI are out there?
- What can I create with generative AI?

Methode: Library research, Literature study.
Date: 09-02-2023

Introduction

For my graduation project I will investigate how generative AI can be used within GroupM. The reason for this being that AI has the potential to automate progresses within GroupM easier and efficiently. It has lots of business potential and is for this reason important to look into possible future applications.

But before diving deeper into the subject I found it important to get a better understanding of what generative AI / Synthetic Media is, as I am new to the subject. By knowing the new technology I will use for my assignment will result in a better outcome of possible products. In this research I will be looking at the history of generative AI, existing tools and how they work.

Generative AI

Artificial Intelligence (AI) is a computer system that is able to perform tasks which normally need human intelligence. With the help of machine learning it has the ability to learn without it being programmed.

Generative AI is a term used to describe an AI that mainly focuses on creating new content based on existing data. Generative AI can recognize patterns and make predictions based on your input. The created content by this AI is called synthetic media.

I asked the trending ChatGPT AI to explain what generative AI is. It answers as follows:

Generative AI refers to a subset of Artificial Intelligence (AI) that involves the creation of new data based on existing patterns and relationships. The goal of generative AI is to generate new content, such as images, videos, music, and text, that resembles existing data.

This is achieved through the use of generative algorithms, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), which are trained on large datasets to recognize patterns and relationships in the data. The algorithms then use this knowledge to generate new content that mimics the style and characteristics of the original data. (ChatGPT, 2023)

In short, generative AI creates content that is not made by humans.

History

Generative AI is becoming popular by the day. For this reason it is important to understand its roots. Early records of generative AI can be found in the 1930s and 1940s when the concept of artificial intelligence that could generate random outputs was explored by researchers. It was also when early pioneers which included Alan Turing, a mathematician started working on early basic techniques of machine learning. In the 1970s they developed the first powerful computers to run them. (McKinsey & Company, 2023)

Generative AI wasn't created by just one person. But through the advancement of neural networks and machine learning in the 1990s and 2000s the generative AIs got more sophisticated. The generative models could create realistic images, audio and video.

An important development in 2014 was the creation of Generative Adversarial Network (GAN) by Ian Goodfellow and colleagues. GANs are an algorithm that are trained using a two-part model. One named a generator, this will generate synthetic data. The second named discriminator learns what is real and synthetic. It will eventually learn how to produce output that is realistic. (Google Developers, 2023)
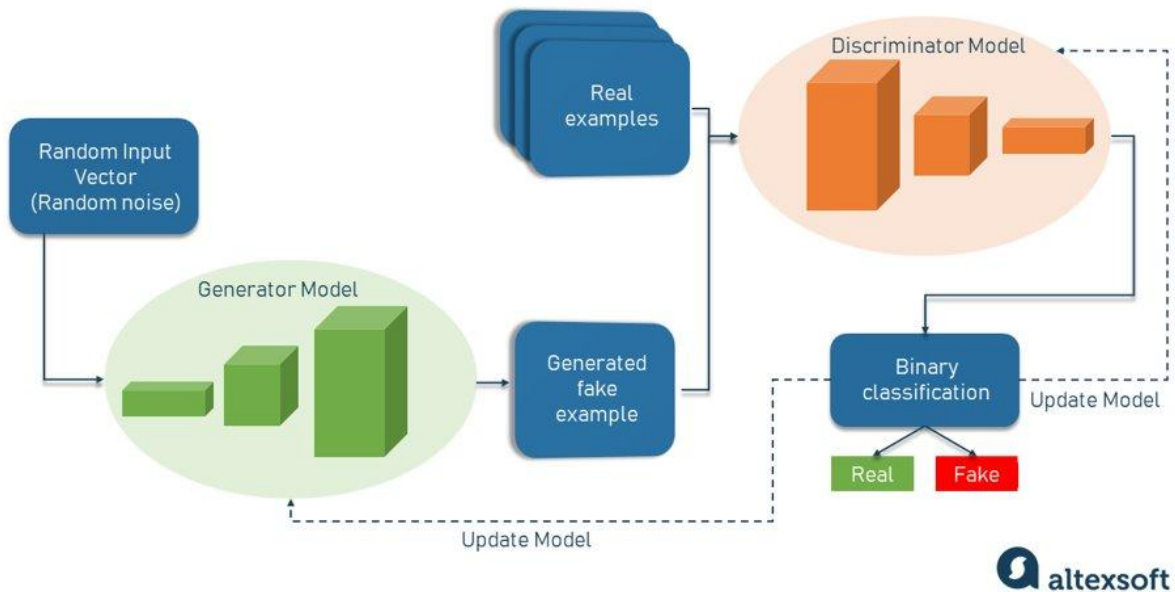
Figure 1: Generative Adversarial Network (GAN) architecture

Since this big advancements in the generative AI field, researchers are exploring more possible options for generative AI.

Different Generative AI

Generative AI has lots of different synthetic media it can create. It can range from simple text or image to video and audio. It usually starts off with a prompt. A prompt is a short description with words and phrases. The newer generative models combine more than one capability. They are called multimodal models, they understand several types of content and let them convert one into the other. Sometimes even transforming them into a completely different content. (Wurmser, 2023) ([Link](#)) The content that generative AI now can generate involve:

- Text
- Images
- Sound
- Speech
- Music
- Video
- 3D
- Animation

Each form of synthetic media could be used for different purposes. There are a lot of possibilities when using generative AI. Some are more developed than others.
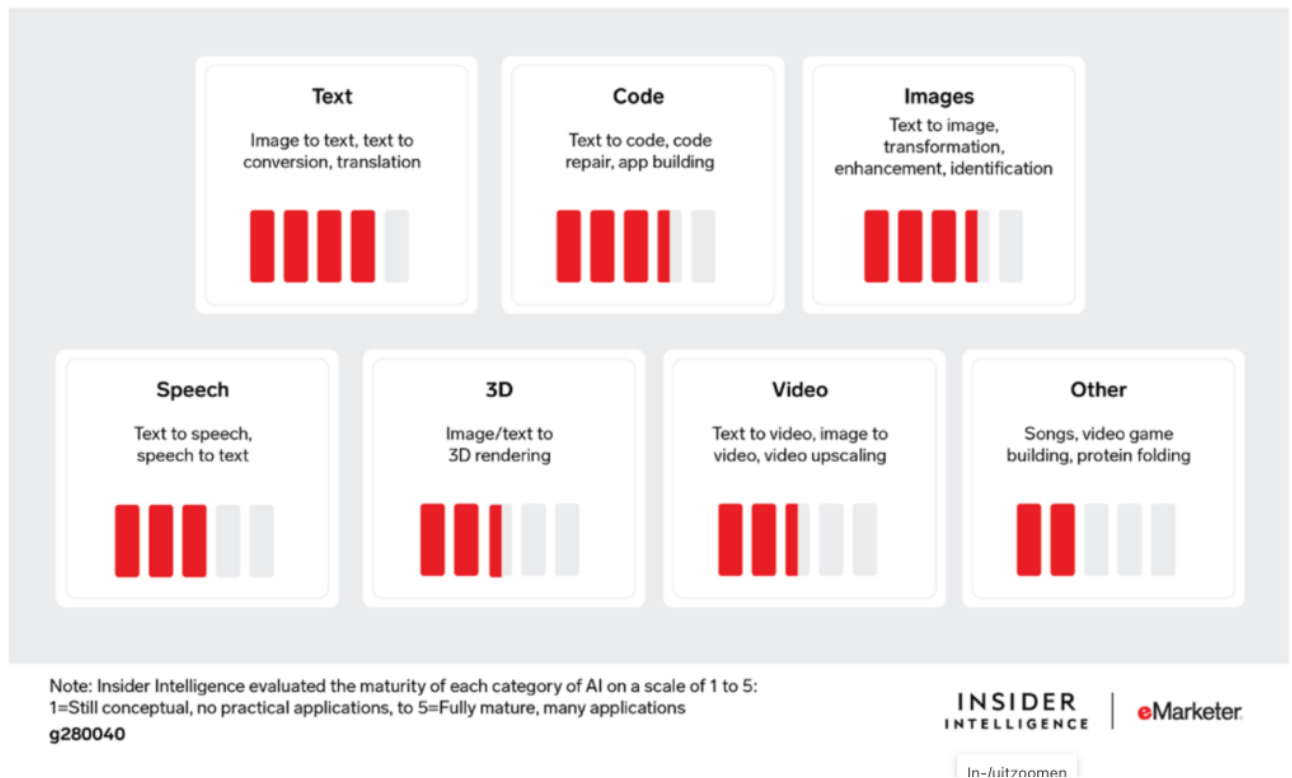
Maturity of Generative AI, by Category

Text	Code	Images
Image to text, text to conversion, translation	Text to code, code repair, app building	Text to image, transformation, enhancement, identification

Speech	3D	Video	Other
Text to speech, speech to text	Image/text to 3D rendering	Text to video, image to video, video upscaling	Songs, video game building, protein folding

Note: Insider Intelligence evaluated the maturity of each category of AI on a scale of 1 to 5: 1=Still conceptual, no practical applications, to 5=Fully mature, many applications
g280040

INSIDER INTELLIGENCE | eMarketer

In-/uitzoomen

Figure 2: Maturity of Generative AI (Wurmser, 2023)

As we see in figure 2 that on top of the most developed generative AI is Text. Coding AI's and image generators are also well developed but all different AI's are rapidly advancing. There are many different platforms for working with AI. You have online web-based AI's, cloud for example Microsoft Azure, mobile apps, etc. The most commonly used by the public is web-based generative AI's. For these AI's you don't have to install software and can easily experiment with these technologies. For my own experience I will also test out some generative AI's.

ChatGPT

ChatGPT is one of the most trending generative AI's. It is a conversational AI that gives information based on request or question. ChatGPT stands for generative pretrained transformer and is built by OpenAI. It was released November 30, 2022 as a proof of concept. In a short span it gained 1 million users in just 5 days. It is text to text.

what is chatgpt

GPT-3 (Generative Pretrained Transformer 3) is a language model developed by OpenAI. It is a deep learning-based natural language processing (NLP) model that has been trained on a massive amount of text data from the internet. As a result, it has a large and diverse knowledge base that enables it to generate text that is coherent and consistent with a given prompt.

ChatGPT is a specific instance of GPT-3 that has been fine-tuned for the task of conversational AI. This means that it has been trained on a smaller, more specific dataset that focuses on generating human-like responses in a conversational setting. With its large knowledge base and ability to generate coherent text, ChatGPT can be used to build conversational interfaces, such as chatbots, that can understand and respond to user input in a human-like way.

The next question I asked is: What can I do with ChatGPT?

ChatGPT can be used for a variety of tasks in the field of natural language processing (NLP), including:

1. Chatbots: ChatGPT can be used to build chatbots that can respond to user input in a natural and conversational way. These chatbots can be used in a variety of applications, such as customer service, information retrieval, and entertainment.
2. Text generation: ChatGPT can be used to generate text based on a given prompt. This can be used in a variety of applications, such as creative writing, content generation, and language translation.
3. Question answering: ChatGPT can be used to answer questions based on its large knowledge base. This can be useful for information retrieval and as a research tool.
4. Text classification and sentiment analysis: ChatGPT can be used to classify text into different categories, such as spam or genuine, or to analyze the sentiment expressed in a given text.
5. Text completion and suggestion: ChatGPT can be used to complete a partially written sentence or suggest the next word in a sentence, based on the context.

These are just a few examples of what ChatGPT can be used for. As NLP technology continues to advance, there is likely to be a growing range of applications for models like ChatGPT in a variety of fields.

ChatGPT is a chatbot where you can ask questions and it will generate an answer. There are many uses for ChatGPT and it works really well. These are some of the reasons why it is popular.

Problems

ChatGPT is very useful but it is not perfect.

- Generative AI models have biases as they are trained on limited data sets.
- ChatGPT needs vast computing power which results in every query costing a few cents.
- Lots of bugs such as that ChatGPT runs only on data provided in 2021, Accuracy it can provide false or incomplete information, inconsistency it loses functionality as OpenAI removes functions that work poorly.
- Limitations on search queries. (Wurmser, 2023)
- You have no insight in their data set so it is questionable how this generative ai is trained.

DALL-E

DALL·E 2 is a new AI system that can create realistic images and art from a description in natural language. (OpenAI, 2022) It is also created by OpenAI. DALL-E was revealed by OpenAI in a blog post on January 5, 2021 and it was opened for the public on 30 September 2022. (NPR Staff, 2022)

DALL-E uses prompts to create digital images. You can go into much detail while writing your prompt resulting in different images. When writing a text prompt you can make adjustments and create images. The generative AI works from text to image.

**Prompt:** future realistic world city neon lights

Prompt: future realistic world city neon lights lightning painting
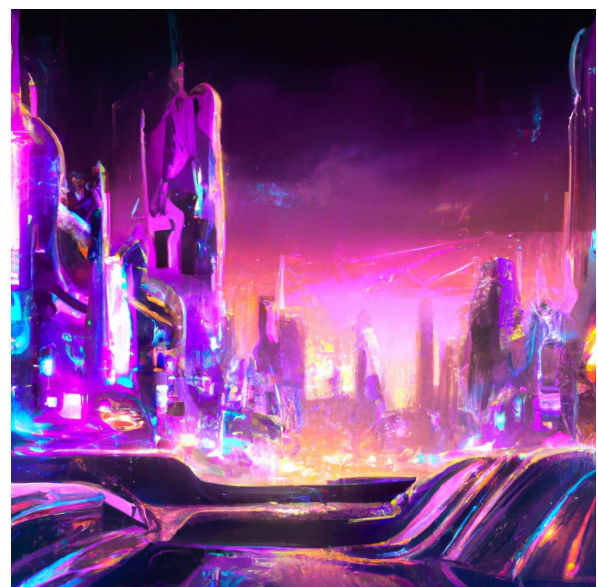
As you see above the generated images look really cool and are adjustable by prompt. It is easy and quickly visualized. Some of the uses for these image generators could be visualizing concepts in early stages, inspiration gathering, storytelling, etc.

Stable diffusion

Stable diffusion is just like DALL-E is a generative AI that creates images. It works the same as DALL-E starting with a prompt that will turn into an image. So text to image. Stable Diffusion gives me four possible images that I would like to see. A few more cartoonish and one highly realistic.
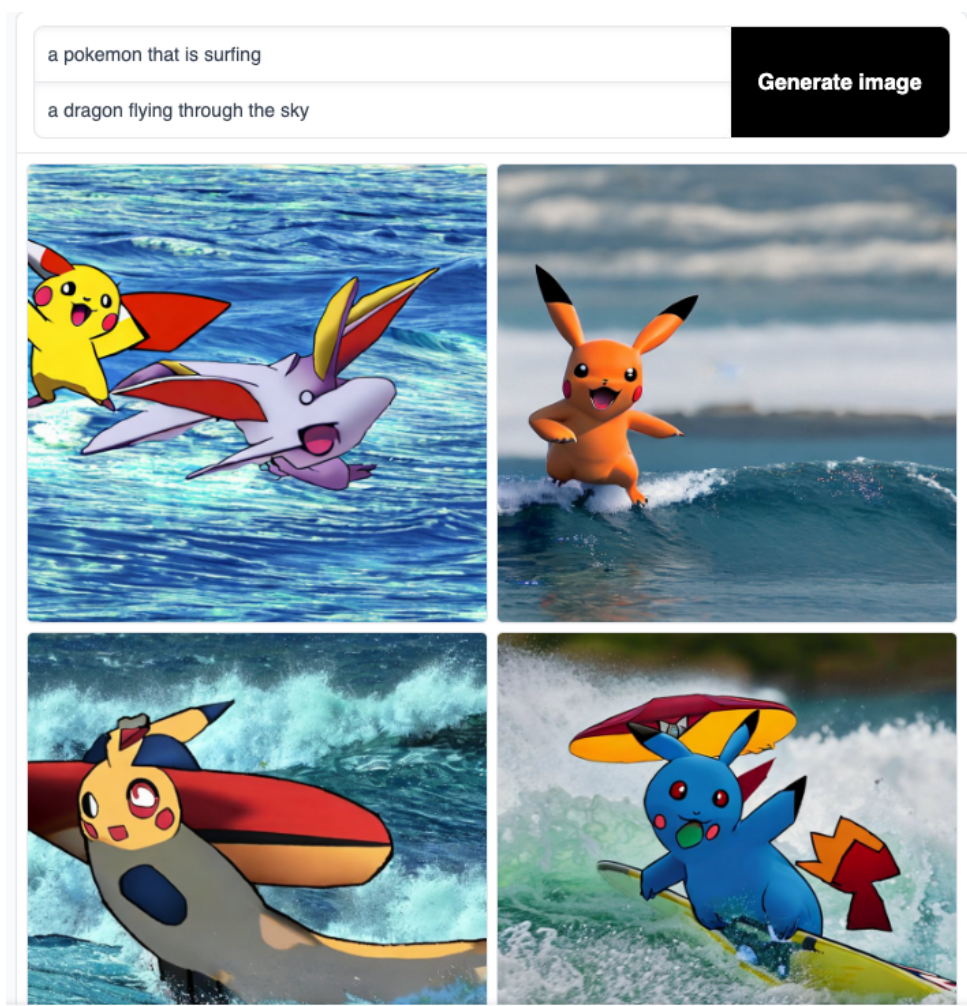
Problems

Possible problems with these image generators could be:

- Lack of diversity: Generated images are trained on a data set and can often be limited. An example of this is that most of the time these generate humans as white and not often as other ethnicities.
- Quality: The quality of the generated images can be low, with visible artifacts and blurriness.

- Anatomy: Often image generators have difficulty with realistic anatomy of humans or other organisms which makes the image instantly not realistic. For example, more fingers than five.
- Copyrighted: Some data sets are trained on copyrighted images and can fall in bad taste to the actual content creators.

Conclusion

Generative AI is a term to describe an AI that mainly focuses on creating new content based inspired researchers to search for new possibilities with these generative AI models. Which leaves us now being that generative AI can create highly realistic images, helps with questions, creating videos, audio, etc.

How to create content is simple, most of the time you start with a text prompt and it will generate new content from this prompt. These are the more traditional generative AI. The more sophisticated use multiple inputs to generate an output. The best working AI is text to text but there are also a lot of other possibilities than only text. There also is video to video, image to image, speech to video, etc.

Although generative AI can be used for many possibilities there are some limitations and problems with these technologies that one should be aware of. One legal problem being that there is no insight on which data the generative AI is trained which means copyrighted content could be used.

To get a better understanding of what impact generative AI has in the creative and marketing sector I will do research on the trend. With the insights of this research I will understand the urgency, relevance of doing research in generative AI for GroupM.

Literature List

Background: What is a Generative Model? (2023). Google Developers. Geraadpleegd op 9 februari 2023, van https://developers.google.com/machine-learning/gan/generative

ChatGPT. (2023). Open.ai. Geraadpleegd op 9 februari 2023, van https://openai.com/blog/chatgpt/

Goodfellow, I. J. (2014, 10 juni). Generative Adversarial Networks. arXiv.org. https://arxiv.org/abs/1406.2661

NPR Staff. (2022, 30 september). DALL-E is now available to all. NPR put it to work. npr.org. Geraadpleegd op 13 februari 2023, van https://www.npr.org/2022/09/30/1125976146/dall-e-art-ai-generator-npr

OpenAI. (2022, 14 april). DALL·E 2. OpenAI. Geraadpleegd op 13 februari 2023, van https://openai.com/dall-e-2/

Overview of GAN Structure. (2023). Google Developers. Geraadpleegd op 9 februari 2023, van https://developers.google.com/machine-learning/gan/gan_structure

What is generative AI? (2023, 19 januari). McKinsey & Company. Geraadpleegd op 9 februari 2023, van https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai

Wurmser, Y. (2023, februari). ChatGPT and Generative AI. Insider Intelligence, van https://www.insiderintelligence.com/content/chatgpt-generative-ai