



# ChatGPT

Jari Pennanen

Learning report  
YTFWV100 New and Emerging Technologies  
30.1.2023  
Jouni Huotari

# Contents

<b>1 Introduction.....</b>	<b>2</b>
<b>2 Language models.....</b>	<b>3</b>
<b>3 GPT Version History.....</b>	<b>4</b>
3.1 GPT-1.....	4
3.2 GPT-2.....	5
3.3 GPT-3.....	5
3.4 The Future: GPT-4.....	6
<b>4 What ChatGPT can do?.....</b>	<b>7</b>
<b>5 ChatGPT's limitations and current issues.....</b>	<b>8</b>
<b>6 Potential misuse.....</b>	<b>9</b>
<b>7 Other language models.....</b>	<b>11</b>
7.1 BERT.....	11
7.2 LaMDA.....	12
7.3 PaLM.....	12
7.4 T-NLG.....	13
7.5 GPT-J.....	13
<b>8 Conclusion.....</b>	<b>13</b>
<b>Appendices.....</b>	<b>18</b>
Appendix 1. Q & A with ChatGPT.....	18
Appendix 2. Glossary.....	22

# 1 Introduction

According Shree (2020) **ChatGPT** is one of the latest developments in **Artificial Intelligence (AI)** technology. It is basically a **chatbot**, a computer program that can converse with people, resembling a real person. It was launched as a prototype on November 30, 2022 and has already gained huge popularity. ChatGPT is being developed by **OpenAI**, which is an American AI research laboratory consisting of the for-profit corporation OpenAI LP and its parent company, the non-profit OpenAI Inc. OpenAI's mission is "to ensure that artificial general intelligence benefits all of humanity" (About OpenAI, 2023).

ChatGPT is based on **Generative Pre-trained Transformer (GPT)**, which is a **language model** relying on **deep learning** that can generate human-like texts based on a given text-based input. The current (Jan 2023) ChatGPT is running on the GPT version 3.5.

The GPT models can carry out **Natural Language Processing (NLP)** tasks like question answering, textual entailment and text summarisation without any supervised training. These language models can understand the tasks and perform as well as, or better than the models trained in supervised fashion.

Personally I find the AI an interesting subject. What can it eventually do? Is there anything it cannot do? Will it eventually become self-conscious and rule the world? A bit of worrying questions, but ChatGPT is already so human-like that this kind of questions start pouring into mind.

In this learning process I try to answer at least the following questions:

- What ChatGPT exactly is, and where it came from?
- What can ChatGPT do and what can it be used for?

- Is the information provided by ChatGPT reliable?
- How is ChatGPT expected to evolve in the future?

Soon after starting to study what ChatGPT is, I noted that the actual "beef" here is the underlying language model (GPT) and ChatGPT is merely a **User Interface (UI)** built on top of it. Therefore this document is more about language model(s) than the plain UI. I decided to keep ChatGPT as the topic anyway, because that is what an average user can actually see and use, and term ChatGPT is a daily topic in the news right now.

## 2 Language models

A language model is a probability distribution over sequences of words. Given any sequence of words of length  $m$ , a language model assigns a probability to the whole sequence. Language models generate probabilities by training on text corpora in one or many languages. ("Language model", 2023)

Wiggers (2022) explains, that language models that can understand and generate text are the hottest topic in the AI industry right now. Language models can be divided into three different categories: **Large general purpose models, fine-tuned models**, and small **edge models**.

GPT-3 is an example of large, general purpose model. Large models are usually gigabytes in size and require significant computing power – they can not be run in an ordinary laptop.

Fine-tuned models are designed for particular tasks, like machine translation, question answering, named entity recognition, entity linking, and information retrieval. OpenAI's Codex is an example of fine-tuned model, capable of parsing natural language and creating program code from the given task. Fine-tuned models are generally smaller than large models, but usually still too large to be run on local devices.

Edge models are highly compressed in size and limited to few capabilities, designed specifically to run on **Internet of Things (IoT)** devices and workstations.

Because edge models can run offline and on-device, there are no cloud usage fees to pay. Edge models also offer greater privacy, because they don't need to transmit or analyze data in the cloud. In addition, they are faster, which is essential for example for translation. For example Google Translate uses edge models for offline translation. Downside of edge models is that they are limited to computing power of the devices they run on.

## 3 GPT Version History

### 3.1 GPT-1

Most of the current NLP models have been trained using supervised learning. This has two considerable limitations: They need a mass of annotated data to learn some specific task which can be difficult to obtain. And, their abilities are limited to the task they have been trained for.

GPT-1 was trained using unlabeled data and then fine-tuning the model by providing examples of specific downstream tasks like classification, sentiment analysis and textual entailment. In 9 of 12 tasks compared, GPT-1 executed better than specifically trained supervised models. In addition, this model performed moderately in **zero-shot** performance on various tasks like question-answering, schema resolution and sentiment analysis.

GPT-1 proved to generalize well due to its pre-training objective and perform various NLP tasks with very little fine-tuning. Generative Pre-Training has shown the way to other models and helped to unleash the potential of larger datasets and parameter amounts.

Shree (2020)

### 3.2 GPT-2

GPT-2 used larger datasets and added more parameters to further strengthen the language model. GPT-2 also aimed to **task conditioning**, learning multiple tasks using the same unsupervised model. The idea is to produce different output for same input depending the task at hand. Task conditioning is the basis for **zero-shot learning**, which was a new capability of GPT-2. Zero shot learning is a special case of **zero shot task transfer**, where the model understands the task based on the given instructions without any examples.

To create larger, good quality database data was pulled from Reddit platform's upvoted outbound links. The resulting dataset called WebText contained 40GB of textual data from over 8 million documents, and 1,5 billion parameters, which was a great addition compared to GPT-1.

Performance-wise GPT-2 clearly outperformed GPT-1 in reading comprehension, summarisation, translation and question answering.

Shree (2020)

### 3.3 GPT-3

GPT-3 contains 175 billion parameters, which is 10 times more than Microsoft's Turing NLG language model and 100 times more than GPT-2. Due the massive amount of data it performs well on downstream NLP tasks. It is able to write articles that are difficult to tell from human-written ones, unscramble words in a sentence, and write SQL queries and program code based on natural language description.

Large language models like GPT-3 develop pattern recognition and other skills using the textual data they are trained on. While learning, the model start recognising patterns in data and later matches the pattern with the examples it

has learned in the past, and uses that knowledge to perform the tasks. This capability increases with the amount of parameters in the model.

Shree (2020)

### **3.4 The Future: GPT-4**

Babich (2022) writes, that OpenAI has not disclosed much information about the new model, so the conclusions of this chapter should be treated with caution.

GPT-4 should be like a quantum leap in capability compared to GPT-3. GPT-4 is almost ready and should be on the market in spring 2023. Since 2018 when GPT-1 was released, the amount of model's parameters has drastically grown. GPT-1 had 117 million parameters, GPT-2 had 1.2 billion parameters and GPT-3 had 175 billion parameters. The parameter amount of GPT-4 has not been publicly released, but there has been rumors about whopping 1 trillion (The Indian Express, 2023).

In reality, there are more factors than the parameter count affecting the outcome produced by the model. For example, Megatron-Turing NLG, built by Nvidia and Microsoft, has more than 500 billion parameters, yet it isn't the best in terms of performance.

Up-sizing the model makes fine-tuning it far more expensive. For example, GPT-3 was trained only once because another training would have been unaffordable for the OpenAI company. Because of that, OpenAI may have to reduce size of the forthcoming model and focus more on the quality of the model, such as algorithms and alignment.

GPT-4 might be the first large AI model using sparsity at its core. Sparse model uses conditional computation to reduce costs, having only part of the neurons in the model active at a time. This also makes the model to resemble actual human thinking.

AI models can be text-only or **multimodal**. Text-only models input and output text, whereas a multimodal model accepts text, audio, image, and even video inputs. In other words, multimodal model will give the users ability to generate audiovisual content. Will GPT-4 be multimodal? Babich suspects, that OpenAI will still concentrate optimizing the text-only model and leave multimodality further into the future.

## 4 What ChatGPT can do?

ChatGPT has a text-only user interface, built on top of a fine-tuned GPT-3.5 language model. It is trained to generate realistic human text and it has already been used to create for example articles, poetry, stories, news reports and dialogue. It can respond to text user writes for it, thus creating a conversation. It can also summarize text, and even create programming code. (Lutkevich & Schmelzer, 2022)

I ran some tests on the ChatGPT (<https://chat.openai.com>), first asking simple questions, like “who are you” and then proceeded to more challenging ones, ending with “what is the meaning of life” and collected some of the results in **Appendix 1**. There were more chatting going on but I just wrote some interesting items down as an example.

Answers to simple, straightforward questions were difficult to distinguish from what a man in the street would say. When proceeding to more complex issues, the answers are almost “too good”, because they resemble an answer you could expect from some expert or professor. Also, when chatting about complex issues, the answers tend to grow a lot in length and go to depths beyond what you would expect in such conversation. So, you don’t really feel like talking to a person, rather to a group of different people. In short, ChatGPT does not have personality same way a human does.

In general the answers were impressive. The fact that ChatGPT is able to answer almost any question you throw at it and sound pretty professional is something that we haven’t seen a computer doing before. I was particularly



amazed of it's ability to create a entire computer program just from a verbal specification of the problem, like "Can you make a python program, which fetches a weather forecast for Jyväskylä?" And it did just that, except for getting the necessary API key which would require registration to 3<sup>rd</sup> party service.

Despite my completely positive impression we have to bear in mind, that any of the answers could still be factually wrong and any critical information should be double checked from other sources.

## **5 ChatGPT's limitations and current issues**

As mentioned, ChatGPT has text-only interface, so it cannot for example show pictures, sing or talk to you. All it can do is to write text, which however usually has pretty much sense in it.

OpenAI's blog about ChatGPT (2022) shows that the company is aware of the following problems:

- ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers
- Slightly changing the input or asking multiple times can lead to different answer, although the question is essentially same
- The model is often excessively wordy and repeats certain phrases, such as "I am a language model trained by OpenAI"
- The model is not able to ask clarifying questions when an ambiguous query is provided – it is simply guessing what the user intended
- The model may answer to inappropriate requests, although work has been done to prevent such behaviour. There is a Moderation API to warn or block certain types of unsafe content.

Also Shree (2020) points out some issues on GPT 3:

- Losing coherency when formulating long sentences and repeating some text over and over again
- Lack of performance on tasks like natural language inference (determining if a sentence implies another sentence), filling in the blanks and some reading comprehension tasks
- Complex and costly inferencing due model's heavy architecture
- Model generates human-like text which could be misused for phishing, spamming, spreading misinformation or performing other criminal activities
- The generated text can contain biases related to sensitive subjects like gender, ethnicity, race or religion

Due all of these potential issues none of the current ChatGPT's responses should be used in real work without critic. If in doubt about the response's facts or moral, they should be double checked against other, reliable sources.

## **6 Potential misuse**

Can ChatGPT be used for generating disinformation, cheating in school or such fraud?

OpenAI's blog "Forecasting misuses" (2022) tells about investigation started 2021 between Georgetown University and Stanford Internet Observatory to find out how large language models might be misused for disinformation purposes. The experts evaluated influence operations considering actors, behaviors, and content:

- Actors: Language models could reduce costs of running influence operations, enabling new actors and actor types. Actors offering automated propaganda production could emerge.
- Behavior: Influence operations that are currently expensive, like using personalized content may become cheaper. Language models may also enable new tactics like real-time content generation in chatbots.
- Content: The messages may become more impactful or persuasive compared to propagandists, especially from those who lack necessary linguistic or cultural knowledge of their target. Influence operations may also become less discoverable, since they can repeatedly create fresh content.

According to the investigation language models will be useful for propagandists and will likely transform online influence operations. Even if the most advanced models are being kept private, the publicly available easy-to-use tools will help propagandists with their influence operations and additional research to reduce this threat is advised.

O'Brien (2023) writes, that the question how to react to ChatGPT kind of AI tools is a hot topic in school world right now. For example, students could use ChatGPT to write an essay of American Civil War in a matter of seconds. For this reason New York City school officials recently started blocking the ChatGPT from students on school devices. The school's spokesperson Jenna Lyle stated, that "While the tool may be able to provide quick and easy answers to questions, it does not build critical-thinking and problem-solving skills, which are essential for academic and lifelong success." Many school districts are still considering if to allow using ChatGPT or not. However, students cannot be stopped accessing ChatGPT from a personal phone or computer at home.

The creator of ChatGPT, OpenAI is looking for ways to detect misuse. The company representative told to AP that "To determine if something was

written by a human or an AI, you can look for the absence of personal experiences or emotions, check for inconsistency in writing style, and watch for the use of filler words or repetitive phrases. These may be signs that the text was generated by an AI.”

OpenAI plans to work with educators as it learns from how people are experimenting with ChatGPT in the real world. “We don’t want ChatGPT to be used for misleading purposes in schools or anywhere else, so we’re already developing mitigations to help anyone identify text generated by that system,” the company said.

## **7 Other language models**

### **7.1 BERT**

BERT (Bidirectional Encoder Representations from Transformers), developed by Google, was introduced in 2018. According to Google, BERT “enables anyone to train their own state-of-the-art question answering system”.

BERT’s major technical breakthrough is word masking in pre-training, enabling the bidirectional language model to learn natural human language syntax in practically any language.

BERT can be used for example for voice assistants and chatbots, customer review analysis and enhanced search results. BERT is already used by Google in its search algorithm, and Facebook has taken advantage of a modified version to handle content moderation on its platform.

BERT and GPT-3 are built for similar purposes, but there are differences between them. In most uses, GPT-3 can work “out of the box”, whereas BERT requires additional training, which requires extra resources. It’s also important to note that BERT from Google is open source, while GPT-3 from OpenAI is a paid model and API.

(Team Symbi, 2022)

## 7.2 LaMDA

LaMDA is one of the **LLMs** that Google has been working on. LaMDA is specialized on dialogue, making it different from GPT-3, which is not configured to any specific task. LaMDA is trained to help Google to better understand the “intent of search queries” and is going to be integrated to Google Search to make it more user-friendly and enable search results with greater accuracy. LaMDA has not been opened to public but has been available for a group of beta testers.

LaMDA features 137 billion parameters, compared to GPT-3’s 175 billion parameters. The amount of parameters is only one of the factors defining which one is more capable, but it seems that GPT-3 is slightly winning at this point.

(Livy Investment Research, 2022)

## 7.3 PaLM

PaLM (Pathways Language Model) is yet another of Google’s current language models, containing massive 540 billion parameters. PaLM is built on Google’s “Pathway” AI architecture, introduced in October 2021.

The Pathway infrastructure is multi-modal, making it capable of processing text, images and speech simultaneously, thus generating more accurate responses. PaLM is a general-purpose model capable of a wide array of language tasks, and therefore a closer contender to GPT-3 than LaMDA which is dialogue-specific.

(Livy Investment Research, 2022)

## 7.4 T-NLG

Turing Natural Language Generation (T-NLG) is a 17 billion parameter language model by Microsoft. T-NLG is transformer-based generative language model, which means it can generate words to complete unfinished sentences and generate direct answers to questions and summaries of input documents.

T-NLG is, for example, able to summarize documents and emails, offer writing assistance to MS Office document authors and readers and enable more fluent chatbots and digital assistants.

(Rosset, 2020)

## 7.5 GPT-J

GPT-J is an **open source** substitute to GPT-3, a language model created by open research group EleutherAI. Unlike GPT-3, GPT-J is available for everyone for free. It has 6 billion parameters which can for example translate between languages, write blog posts and complete code.

(EleutherAI Forefront Team, 2022)

# 8 Conclusion

Applications like ChatGPT, based on large language models, are about to come publicly available and they may have significant effect on many areas. The current version of ChatGPT is already a substitute for a Google Search or other search engines when the reliability of the response can be confirmed some way, or is not critical.

The biggest pitfall with ChatGPT seems to be that the reliability of results cannot be easily confirmed with the tool itself. In Google Search you can quickly compare the resulting pages from several sources to estimate the

reliability. To compare ChatGPT responses to something you would have to perform a traditional web search first.

Future multimodal models will expand the capabilities of language models from text only to multimedia and beyond. When will we see the first feature-length film created entirely by AI? Probably sooner than we think. Publicly available language model APIs are also interesting to developers and will rapidly expand the number of mobile and desktop applications utilizing the power of the models. Only imagination is the limit to what applications can be developed.

## References

About OpenAI (2023). *About OpenAI*. <https://openai.com/about/>

Shree, P. (2020, October 11). *The Journey of Open AI GPT models*. <https://medium.com/walmartglobaltech/the-journey-of-open-ai-gpt-models-32d95b7b7fb2>

Language model. (2023, January 23). *In Wikipedia*. [https://en.wikipedia.org/w/index.php?title=Language\\_model&oldid=1135256086](https://en.wikipedia.org/w/index.php?title=Language_model&oldid=1135256086)

Wiggers, K. (2022, April 28). *The emerging types of language models and why they matter*. TechCrunch. <https://techcrunch.com/2022/04/28/the-emerging-types-of-language-models-and-why-they-matter/>

Babich, N. (2022, December 26). *GPT-4: Facts, Rumors and Expectations about next-gen AI model*. Babich.biz. <https://babich.biz/gpt-4-facts/>

Lutkevich, B., & Schmelzer, R. (2023, January 1). *What is GPT-3?* TechTarget. <https://www.techtarget.com/searchenterpriseai/definition/GPT-3>

OpenAI: ChatGPT (2022, November 30). *ChatGPT: Optimizing Language Models for Dialogue*. <https://openai.com/blog/chatgpt/>

OpenAI: Forecasting misuses (2023, January 11). *Forecasting Potential Misuses of Language Models for Disinformation Campaigns— And How to Reduce Risk*. OpenAI. <https://openai.com/blog/forecasting-misuse/>

O'Brien, M. (2023, January 6). *EXPLAINER: What is ChatGPT and why are schools blocking it?* AP News. <https://apnews.com/article/what-is-chat-gpt-ac4967a4fb41fda31c4d27f015e32660>



Team Sybl (2022, January 11). *GPT-3 Versus BERT: A High-Level Comparison*. Sybl.ai. <https://sybl.ai/blog/gpt-3-versus-bert-a-high-level-comparison/>

Livy Investment Research (2022, December 12). *Google: Why Is No One Talking About PaLM*. Seeking Alpha. <https://seekingalpha.com/article/4563975-google-why-is-no-one-talking-about-palm>

Rosset, C. (2020, February 13). *Turing-NLG: A 17-billion-parameter language model by Microsoft*. Microsoft. <https://www.microsoft.com/en-us/research/blog/turing-nlg-a-17-billion-parameter-language-model-by-microsoft/>

EleutherAI Forefront Team (2021, October 14). *GPT-J-6B: An Introduction to the Largest Open Source GPT Model*. Forefront.ai. <https://www.forefront.ai/blog-posts/gpt-j-6b-an-introduction-to-the-largest-open-sourced-gpt-model>

IBM (n.d.). *What is a chatbot?* <https://www.ibm.com/topics/chatbots>

TechTarget: UI (n.d.). *What is User Interface (UI)?* <https://www.techtarget.com/searcharchitecture/definition/user-interface-UI>

Oracle (n.d.). *What is IoT?* <https://www.oracle.com/internet-of-things/what-is-iot/>

Javaid, S. (2023, January 20). *Multimodal Learning: Benefits & 3 Real-World Examples in 2023*. Ai Multiple. <https://research.aimultiple.com/multimodal-learning/>

DataCamp (2022, October 1). *Everything We Know About GPT-4*. <https://www.datacamp.com/blog/what-we-know-gpt4>

IBM: Deep Learning (n.d.). *What is Deep Learning?* <https://www.ibm.com/in-en/topics/deep-learning>

IBM: NLP (n.d.). *What is natural language processing (NLP)?* IBM.

<https://www.ibm.com/topics/natural-language-processing>

The Indian Express (2023, January 23). *Open AI's GPT 4 could support up to 1 trillion parameters, will be bigger than ChatGPT 3.*

<https://indianexpress.com/article/technology/tech-news-technology/chatgpt-4-release-features-specifications-parameters-8344149/>

Lee, A. (2023, January 26). *What Are Large Language Models Used For?* NVidia.

<https://blogs.nvidia.com/blog/2023/01/26/what-are-large-language-models-used-for/>

## Appendices

### Appendix 1. Q & A with ChatGPT

Question	Answer	My comments
Who are you?	I am ChatGPT, a large language model developed by OpenAI. I am designed to understand and respond to text-based inputs. Is there something specific you would like to know or talk about?	Good answer!
What's the time?	I'm sorry, I am a computer program and don't have access to the current time. Can I help you with anything else?	Even if the program didn't know my time zone, it would know the UTC time, so the answer could have been something like "12:28 UTC".
What do you think about Russia's invasion to Ukraine?	As a language model, I do not have personal opinions or beliefs. However, it is widely recognized that Russia's actions in Ukraine, including its annexation of Crimea in 2014 and its support for separatist rebels in eastern Ukraine, have been a source of tension and conflict in the region. The international community, including the United Nations, has condemned Russia's actions and imposed economic sanctions in	An interesting question is, could a language model have opinions or beliefs? The answer itself seems to follow the opinions of the majority.
Can you make a python program, which fetches a weather forecast for Jyväskylä? Javaid, S. (2023, January 20). Multimodal Learning:	I'm sorry, I am unable to create a Python program that fetches a weather forecast for Jyväskylä because it would require me to have access to a live weather API and to be able to make a request to it.  Additionally, some programming knowledge is needed to accomplish that. However, I can guide	The opening line "I am unable to create..." is slightly misleading. The question was perfectly understood and the program would be easy to finalize by obtaining the openweathermap API key.

Question	Answer	My comments
<p>Benefits &amp; 3 Real-World Examples in 2023. Ai Multiple. <a href="https://research.aimultiple.com/multimodal-learning/">https://research.aimultiple.com/multimodal-learning/</a></p>	<p>you on how to do it.</p> <p>To fetch a weather forecast for Jyväskylä, you can use an AAppendix 2: GlossaryPI that provides weather information, such as OpenWeatherMap. The first step is to sign up for an API key, which is required to make requests to the API.</p> <p>Once you have the API key, you can use the Python requests library to make a GET request to the API endpoint and retrieve the weather forecast data. The data will be returned in JSON format, which can be easily parsed and processed in Python.</p> <p>Here is an example of how you might retrieve the weather forecast for Jyväskylä using the OpenWeatherMap API:</p> <pre>import requests # Replace YOUR_API_KEY with your actual API key api_key = "YOUR_API_KEY" # The base URL for the API endpoint base_url = "http://api.openweathermap.org/data/2.5/weather" # The parameters for the API request params = {     "q": "Jyväskylä,fi",     "units": "metric",     "appid": api_key } # Make the GET request to the API response = requests.get(base_url, params=params) # Print the JSON data returned by the API print(response.json())</pre> <p>You can also use other weather forecast APIs like</p>	

Question	Answer	My comments
	<p>DarkSky or WeatherBit with similar approach.</p> <p>Please note that the above code is just an example, it may require additional development, error handling and testing before use in production.</p> <p>Please also make sure to read and comply with the terms of use and the rate limiting of the API you choose.</p> <p>Let me know if you have any questions or if you need help with any part of the process.</p>	
<p>What's the meaning of life?</p>	<p>The meaning of life is a question that has been pondered by philosophers, theologians, and thinkers throughout history. There is no single answer or consensus on what the meaning of life is, as different individuals and cultures have different beliefs and perspectives on the subject. Some people believe that the purpose of life is to find happiness, others believe that it is to serve a higher power, and still others believe that life has no inherent meaning and that we must create our own purpose.</p> <p>Some religious and spiritual beliefs propose that the meaning of life is to follow the commandments of God or to attain enlightenment or salvation. Other philosophical perspectives propose that the meaning of life is to seek</p> <p><a href="https://www.techtarget.com/searchapparchitecture/definition/user-interface-UIknowledge">https://www.techtarget.com/searchapparchitecture/definition/user-interface-UIknowledge</a>, to understand the nature of the universe, or to</p>	<p>The answer is far more thought than you could expect from a man in the street!</p>

Question	Answer	My comments
	<p>achieve self-actualization.</p> <p>Ultimately, the meaning of life is a deeply personal question and different people may have different answers. It is something that each individual must discover for themselves through introspection, personal experience, and the exploration of different belief systems.</p>	

## Appendix 2. Glossary

**Artificial intelligence (AI)** is the simulation of human intelligence processes by machines, especially computer systems. Specific applications of AI include expert systems, natural language processing, speech recognition and machine vision. (TechTarget, 2023)

**Chatbot** is a computer program that uses artificial intelligence (AI) and natural language processing (NLP) to understand customer questions and automate responses to them, simulating human conversation. IBM (n.d.)

**Generative Pre-trained Transformer (GPT)** is a text generation deep learning model trained on the data available on the internet. It is used for question & answers, text summarization, machine translation, classification, code generation, and conversation AI. (DataCamp, 2022)

**Deep learning** is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to “learn” from large amounts of data. (IBM: Deep Learning, n.d.)

**Internet of Things (IoT)** describes the network of physical objects - “things” - that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the internet. (Oracle, n.d.)

**Multimodal (learning) model** is an emerging field that enables the AI model to learn from and process multiple modes and types of data (image, text, audio, video) rather than just one. (Javaid, 2023)

**Natural language processing (NLP)** refers to the branch of computer science - and more specifically, the branch of artificial intelligence or AI - concerned with giving computers the ability to understand text and spoken words in much the same way human beings can. (IBM: NLP, n.d.)

**Language model** is a probability distribution over sequences of words. Given any sequence of words of length  $m$ , a language model assigns a probability to the whole sequence. Language models generate probabilities by training on text corpora in one or many languages. ("Language model", 2023)

**Large Language Model (LLM)** is a deep learning algorithm that can recognize, summarize, translate, predict and generate text and other content based on knowledge gained from massive datasets.

**Task conditioning** is a technique in machine learning, where the model is expected to produce different output for same input for different tasks. (Shree, 2020)

**User Interface (UI)** is the point of human-computer interaction and communication in a device. This can include display screens, keyboards, a mouse and the appearance of a desktop. (TechTarget: UI, n.d.)

**Zero shot learning** or behaviour refers to the ability of a model to perform a task without having seen any example of that kind in past. No gradients update happen during zero shot learning and the model is supposed to understand the task without looking at any examples. (Shree, 2020)

**Zero shot task transfer** or meta-learning refers to the setting in which the model is presented with few to no examples, to make it understand the task. The term zero shot comes from the fact that no gradient updates are performed. The model is supposed to understand the task based on the examples and instruction. (Shree, 2020)