Leveraging AI for Enhanced Human-Computer Interaction: A Case Study on ChatGPT

By J.H.C. van Gent (2023)

Abstract

This report explores the capabilities of Artificial Intelligence (AI) through the lens of OpenAI's ChatGPT. AI is distinguished from non-AI programs by its iterative learning process, enhancing its capabilities with each interaction. ChatGPT, as a data product, encompasses human-like traits, using advanced algorithms and natural language processing. Its utility spans a lot of sectors, from healthcare to finance, showcasing its ability to automate and refine processes previously requiring human intervention. This report looks into the capabilities of ChatGPT, discussing its role as an interactive tool, its continuous learning through user engagement, and its potential to innovate across various applications. Additionally, it considers the ethical implications of AI deployment in the wake of emerging (European) regulations. The report positively acknowledges ChatGPT's ability to automate tasks by imitating human-like interactions, while also carefully considering necessary safeguards like openness, data protection, and regulation to maintain responsible use of AI.

PART I: Problem Selection, Definition & Motivation, and Human in the Loop

Artificial Intelligence

Artificial Intelligence (AI) is defined as the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. The central problems of AI include such traits as reasoning, knowledge, planning, learning, and communication. These traits are manifested in machines that seek to replicate or simulate human intelligence. (Russell et al. 2009)

When we talk about AI programs, the key understanding can be derived from comparing it with non-AI programs. A non AI program takes in some input and generates some output based on a certain algorithm. The difference with an AI program is that it is not finished after generating the output. It inserts the output as input once again and hereby trains the underlying model (e.g. adjusting the parameters of the algorithm). For example, if an AI playing a game loses after making a certain move, it will use that information to adjust its strategy. This process is graphically displayed below.
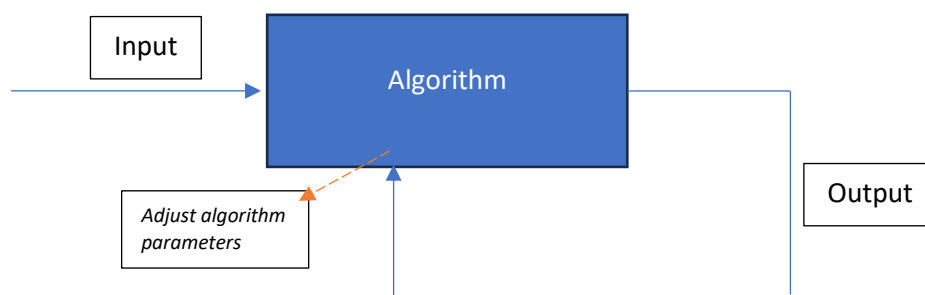

Figure 1: Process conducted by an AI program

Characteristics

To get a clearer sense of what AI is, five important characteristics will be outlined:

Learning and adaptation: Just like humans learn from experience, AI can improve over time. Imagine you're playing a video game and you get better the more you play. AI works similarly. It can learn from data, make adjustments, and get better at whatever task it's doing.

Problem-solving: AI is like a detective in a way. It looks at a problem, considers all the information it has, and tries to find the best solution. For example, it's like an AI playing a game of chess and figuring out the best moves to win the game.

Natural language processing: NLP enables machines to interpret and generate human language, allowing for seamless interaction between humans and computers.

Perception: AI's ability to perceive involves interpreting sensory information to understand and interact with the world. This goes beyond mere image recognition to include the comprehension of complex scenes and auditory signals. Consider autonomous vehicles for example.

Automation: Automation through AI is about enabling systems to perform tasks independently, often in a more efficient and error-free manner than humans. This extends to decision-making processes, where AI can manage and execute tasks without constant human oversight. An example is algorithmic trading. (TheKnowledgeAcademy, z.d.)

PART II: AI Data Product Description of ChatGPT

ChatGPT

ChatGPT is the most popular conversational AI model at the moment. As a data product it combines data science, algorithms, and a textual interface to act as a comprehensive virtual assistant. At it's core, ChatGPT is a sophisticated program that processes natural language input, and generates responses that are relevant and feel human-like. ChatGPT is trained and has several versions. While people use ChatGPT more training is being done.

Problem/Use-Case

Most people have somewhat of an idea what ChatGPT is about. It is a program, which you can chat with to get answers to you questions or to get it to accomplish certain tasks. It is hard to establish one particular problem that ChatGPT solves, and also the use-cases of ChatGPT vary a lot. Some sectors where ChatGPT can be of great use include healthcare, e-commerce, marketing, education, finance, coding, customer service, and many more (George et al, 2023). It is abundantly clear that ChatGPT solves problems in multiple sectors. The common aspect ChatGPT solves is the human interaction that was required in the past.

Capability Domain

ChatGPT is a Natural Language Processing (NLP) program. A few capabilities that describe ChatGPT best:

Contextual Understanding: ChatGPT is smart at getting the hang of conversations. It can figure out what you mean and chat back just like a real person, making talking to it feel more natural.

Language Generation: ChatGPT is really good at writing text that makes sense, stays on topic, and follows the rules of grammar. It's handy for writing articles, summing things up, or changing text to sound different.

Task Adaptability: ChatGPT is like a Swiss Army knife for text tasks. It can be tweaked to do all sorts of things, from helping customers to creating content or teaching stuff. Developers can shape it to fit exactly what they need.

Multilingual Proficiency: ChatGPT can speak many languages, which is great for working with people from all over the world. It's useful for translating, figuring out feelings in text, and creating content in different languages.

Scalability: ChatGPT can grow or shrink to match how much computer power you have or how quick you need it to be. This means it works for both small jobs and big business projects.

(Ray, 2023)

Application Domain

The most basic use case is the implementation of a chatbot. ChatGPT is a chatbot itself. When you access the OpenAI website, you can start asking you questions straight away. But ChatGPT has evolved and more use-cases can be identified. ChatGPT has some great (new) features such as the API, Code Interpreter (now called Advances Data Analysis), and plugins. The API even offers support for function calling. This feature enables developers to give ChatGPT access to custom functions.

Imagine a situation where you have got a leakage at you house. The first thing you do is call a plumber company. Usually some customer support employee picks up the phone and finds an empty

spot in one of the available plumbers calendars. The support employee also needs to estimate how urgent the leakage is. This entire process can be automated by using the ChatGPT API. A developer uses software to get speech to text, and transmits this text through the ChatGPT API. The response from ChatGPT can then be transformed into speech to talk back on the phone. And once ChatGPT has assessed the severity of the leakage it can use a supplied function (via function calling) to get the available slots of the plumbers and book one of these slots.

Data Product Components & Techniques

User Interface

As explained earlier, ChatGPT is not solely the webapp created by OpenAI. Most people view the user interface of ChatGPT to be the well known website, where the user has a textbox at the bottom of the page where it can shoot messages to ChatGPT. This is just the most simple usage.

When we look at ChatGPT as just the webapp, the interface consist out of a few features. The most important part is the conversation display and the input box at the bottom of the screen displayed to the user. When sending a message, the chatbot starts 'typing' it's response. You can choose which AI model to use at the top, and you can view older conversations in the left sidebar. You can also find some new features here (some paid). The beta features at the bottom of the left sidebar can be toggled. As of writing this, two beta features are displayed: plugins, and code interpreter (or Advanced Data Analysis). Furthermore, some basic account and profile settings can be changed.

But the interface of the available ChatGPT functionalities is much larger. An important feature to understand is the API, which is continuously growing in features. The most used API is the Chat Completions API offered by OpenAI. The API let's you define a bunch of parameters, such as the model and the temperature. A conversation usually entails three actors: the system, the user, and the assistant. The system is the application which the developer is working on, and the assistant is the ChatGPT API chatbot. The user is the person using the developer's application.

ChatGPT in terms of its Agency and Architecture

ChatGPT doesn't have consciousness or intent in the way humans do, but it can make decisions and generate responses independently based on its programming and the data it has been trained on. It's agency is entailed in it's ability to understand textual input and generate relevant responses, simulating a form of conversation agency.

As for the architecture of ChatGPT, it is quite complex. It all starts as a Large Language Model (or LLM). An LLM looks like a spider web. It contains multiple layers, where each layer consists of multiple nodes. The nodes point to the nodes in the next layer with weights attached. The GPT-3.5 LLM representation is displayed in figure 3. It has a huge number of parameters (~175 billion) and a lot of layers (~96).
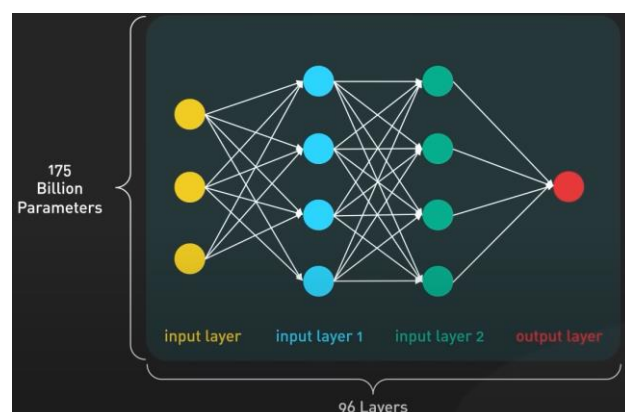
Figure 3: Representation of GPT-3.5 LLM

To represent words, numbered tokens are used, as it works more efficiently (see figure 4). GPT-3.5 used about 500-billion tokens for training.

These tokens are then used for prediction. Given a token or a sequence of tokens, the model will try and predict the next token (or word) (see figure 5).
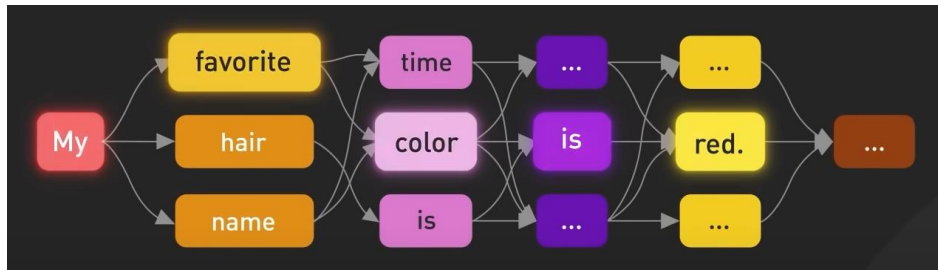

Figure 4: Words as tokens

Figure 5: Predicting a sentence

(ByteByteGo, 2023)

How ChatGPT is trained

The initial training is actually the same as discussed above. The architecture of the LLM is used for the initial training. Words are transformed into tokens (or vectors) and these represent a numerical value. Given a token or a sequence of tokens, the next token is predicted. If the numerical value is off (as the model is controlled), the difference is calculated, and the algorithm is adjusted.

Once the initial learning is finished, the model can be deployed. While in production, the model can keep learning. This is done by analysing user feedback. Consider figure 5 where 'the chef' (represents an AI model like ChatGPT) serves a burger and customers tell him if they liked it or not. If the burger is disliked, it will be removed from the menu and replaced by something else. This is what ChatGPT does with it's chat responses as well.
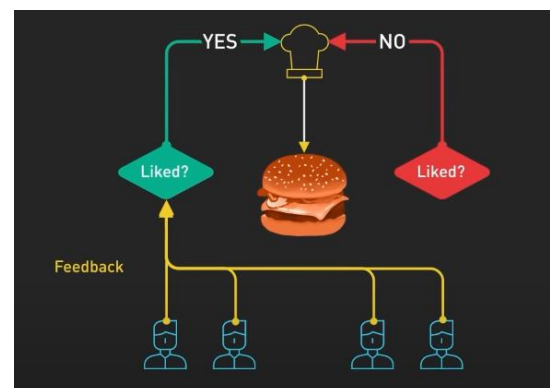

Figure 6: Representation of feedback model

Furthermore, ChatGPT offers a set of parameters (these parameters are different from the parameters in figure 3, which represent nodes in the LLM). A few important parameters are the temperature, the max_length, and the diversity_penalty. These can be used inside ChatGPT to alter it's output. When increasing the temperature, ChatGPT will supply more creative answers. Max_tokens is used to set the amount of tokens to be used by ChatGPT. If you let it have more tokens it can write more extensive responses. And lastly a feature, which is not that often discussed, the diversity_penalty. This parameter can be set to 0, 1, or 2. When applying a higher penalty, the change of getting a 'boring' or 'repetitive' answer declines. (Onslow, 2023)

PART III: Use Case Description & Application

Most popular use-cases

ChatGPT serves as a versatile tool in various scenarios. For writers and content creators, ChatGPT is a source of inspiration and assistance, helping to overcome writer's block and generate creative content ideas. It's like a brainstorming companion that never runs out of suggestions or energy.

In the realm of education, ChatGPT plays the role of a digital tutor, assisting students with homework and explaining complex subjects in a way that's easy to understand. It's akin to a knowledgeable study partner, ready to delve into a wide array of topics at any time.

Additionally, ChatGPT can be employed for language translation, making communication across different languages more accessible. It's like having a multilingual friend who helps you chat with people from around the world.

(Padmavati, 2023)

Examples and Demonstrations

To test ChatGPT in a couple of critical aspects, a few examples will be displayed.

Fact-checking / Truth finding

According to Hoes et al, ChatGPT is capable of conducting fact-checking, but is not super accurate. They tested 21.152 fact-checked statements, and ChatGPT had the write true / false answer in 69% of the cases. This is more than random selection would be, but might not be enough to let ChatGPT handle this task replacing human moderation. (Hoes et al, 2023)

To try it out myself, I found a statement of Trump from back in 2020, which states:

"Frankly, [Biden] ran the H1N1, swine flu, and it was a total disaster, far less lethal, but it was a total disaster. Had that had this kind of numbers, 700,000 people would be dead right now."

The fact-check response of ABC-news here was: "While Trump is correct that the H1N1 virus was much less lethal than COVID-19, it is misleading to call the Obama administration's response a "failure."" (https://abcnews.go.com/Politics/fact-checking-trump-biden-2nd-2020-presidential-debate/story?id=73750341)

When I input the statement in ChatGPT, I get quite an extensive response. It used Bing search and visited several website to come to the following statement:

"*The statement about Biden's handling of the H1N1 swine flu pandemic needs some context for accuracy:*

Joe Biden, as Vice President during the H1N1 pandemic, was part of the administration's response. In April 2009, Biden indeed urged an aggressive early response to the pandemic. However, his public remarks advising against travel on planes or the subway were criticized and contradicted the more measured statements by President Obama, leading to a quick retraction and clarification by other administration officials.

In terms of the impact of the H1N1 pandemic, the Centers for Disease Control and Prevention (CDC) reported that over the course of a year, the H1N1 flu infected 60 million Americans and claimed 12,469 lives in the U.S..

Globally, the pandemic's mortality rate was initially underreported. A study published by the CDC and other collaborators estimated that between 151,700 and 575,400 people died worldwide from the H1N1 virus infection during the first year of its circulation, which is more than 15 times the number of laboratory-confirmed deaths reported to the World Health Organization (WHO).

A notable aspect of the H1N1 pandemic was that it disproportionately affected younger individuals. About 80% of the deaths were in people younger than 65 years, which contrasts with seasonal flu epidemics where 80-90% of deaths occur in those 65 and older.

Despite its significant impact, the global population overall experienced less severe effects from the H1N1 pandemic compared to other pandemics like that of 1968 or the notorious 1918 Spanish flu. The estimated global mortality rate was between 0.001–0.007% of the

world's population for respiratory complications associated with the 2009 H1N1 virus, and slightly higher when including cardiovascular complications.

The claim that 700,000 people would have died if the H1N1 had the same mortality rate as COVID-19 is speculative and does not account for the various factors that influence pandemic responses and outcomes. It is also important to note that the lethality of a virus is just one aspect that affects the overall impact of a pandemic, and the comparison does not consider the differences in transmission dynamics, medical interventions, and public health responses between the two viruses."

It actually did a good job, and it actually used 5 sources (articles) from 2 different respectable entities: politico.com, and the CDC.

Question answering

Question answering is one of the key functionalities of ChatGPT. And does not require demonstration for that reason.

Educational Tutoring

Educational tutoring is actually an aspect that ChatGPT masters really well. By making use of prompt engineering and scripting, great results can be accomplished. A great example is the scripts provided by Mr. Ranedeer, who is an expert in the educational aspect of ChatGPT. When we paste the script v2.7 provided here (https://github.com/JushBJJ/Mr.-Ranedeer-AI-Tutor) into the code interpreter (now Advanced Data Analysis), we will be able to educate ourselves on any topic of out liking.

After pasting the script, you can start by prompting your learning objective, and it will start analysing like so. It generates a curriculum, which you can download, and when hitting the start command, your first lesson will begin.
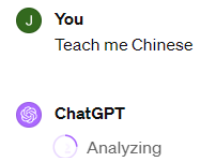
Theory-of-Mind Problem Solving

The theory-of-mind (ToM) problem-solving capacity of ChatGPT is about it's ability to mimic human-like understanding of others' mental states. Researchers have utilized various ToM tasks to evaluate ChatGPT's performance in this aspect. It excels in language-based ToM tasks, suggesting an advanced level of social cognition. However, it's noted that ChatGPT tends to overextend its responses, indicating a divergence from typical human interaction. So it is not completely hum-like. The researchers do state that proper prompting influences this.

Prompt engineering

Prompt engineering is becoming increasingly popular. As shown before in the educational tutoring example, prompting can get very sophisticated. The script of Mr. Ranedeer is a good example of prompt engineering. By leveraging the code interpreter and a well thought out script, he replaces a physical teacher. You are now able to learn whatever subject you'd like.

Prompting engineering is also catching up to well established companies. An employee who knows how to prompt, can work more efficiently, as GPT can write great amounts of code for you. And furthermore, new opportunities can be unlocked with prompt engineering. As earlier explained, levering the GPT API, and the function calling feature, new applications can arise that in the past needed human interactions.

ChatGPT leverages key Human-Computer Interaction (HCI) principles by offering an intuitive conversational interface that allows users to interact with the AI in a natural and human-like manner, enhancing user experience and accessibility. Regarding performance level and accuracy, ChatGPT demonstrates a high degree of proficiency in understanding and generating human language. However, the accuracy can vary based on the complexity of the task and the specificity of the information provided. Good prompt engineering is a key factor here.

PART IV: Critical Reflection & Ethical Considerations

Popularity

If you have never heard of ChatGPT by now, you have probably been living in a cave. Some people even call ChatGPT the most disruptive tech of the century. Figure 7 below really shows it's popularity, as it took ChatGPT just two months to reach 100 million users!
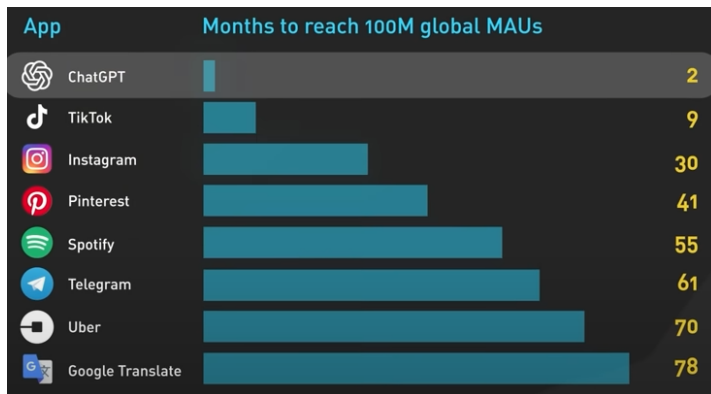

Figure 7: Months to reach 100 million users per application (ByteByteGo, 2023)

Does ChatGPT solve our problem?

ChatGPT certainly does solve our problem, as it enables us to automate tasks that needed a human touch in past. The key characteristics that we spoke about really reflect this. An AI-program is able to 'understand' and act accordingly, whilst a non-AI program is can only follow pre defined instructions.

The use-cases that were discussed show the problem solving ability. They are all tasks that would normally have be carried out by a human, and can now be automated.

The European AI-Act

Also the European Union sees this fast changing and growing environment, and needs to adjust or make laws accordingly to protect it's citizens.

The most important factors discussed in the AI-Act are the prohibition of certain high-risk AI systems. An example given would be an AI powered voice in a kids toy, that could say harmful things. The act also states that systems and algorithms should be transparent and explainable, which is some times hard to do, as of the Blackbox nature the technology brings with it. Another important factor is privacy. Systems should protect personal data and give individuals control over their data. This is also a painful aspect, as models as ChatGPT learn by using user input, which in a sense can harm user privacy. If I were to have a bug in my code and ask ChatGPT to find the problem, whilst forgetting to remove API-keys, it might be possible that ChatGPT is going to predict my API key when generating text in the future. (EU AI Act, 2023)

This was a very brief overview of what is included in the act. A lot of positive thing will come from this. Users will have more trust in the technology, as their data should be protected and harmful AI systems will be banned or thoroughly tested before they can be used in production.

But it also brings potential downsides. The cost and effort of AI systems will increase, as they need to comply with the new regulation. This can lead to a downside to the EU. We've seen with the AI-bot of Google called Bard, that it took quite some time before it became available in Europe, as it did not

yet comply with EU regulations. The new AI-Act can thus lead to a competition disadvantage, and the EU can be seen as a less attractive market for trying new innovations.

Studied Literature

A list of all sources used is displayed below. A couple of relevant sources deserve some extra attention to clarify their usage in the report.

1. *ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. By Ray, 2023.*
   This article does a very thorough job in explaining what ChatGPT is capable of, and what the most common use-cases are. It introduces a few interesting subjects such as Machine Learning, and an explanation on LLM's. It is also critical of the new technology, as it discusses potential safety issues, and ethical concerns. These also clarify some of the points presented in the European AI-Act, as discussed in the report.

2. *Leveraging ChatGPT for Efficient Fact-Checking. By Hoes, E., Altay, S., & Bermeo, J. , 2023.*
   This article is very interesting as it shows the capabilities of ChatGPT to establish the truth. The article states that after supplying 21.152 fact-checked statements, ChatGPT was able to establish correct categorization in 69% of the cases. This is more then a doing it in a random manner (as then it would be closer to 50%), but it is far from water proof. This means it can be concluded that we should not take information provided by ChatGPT as truthful. Although the example provided in this report shows a thorough correct answer on one of my supplied fact-checked statements. This gives the impression that ChatGPT is improving.

3. *A review of ChatGPT AI's impact on several business sectors. Zenodo (CERN European Organization for Nuclear Research). George, A., George, A., & Martin, A., 2023.*
   The article gives some background information on the establishment of the current GPT-versions, and gives an overview of use-cases in different sectors. It is of great relevance for the report as it actually shows that ChatGPT is actually solving a similar problem in these different sectors. ChatGPT is able to automate some actions, that normally require human intervention. ChatGPT is even able to mimic human-like understanding of other's metal states. In this report we do state that this might require some prompt engineering to correctly accomplish.

APA Reference List

ByteByteGo. (2023, 24 april). How ChatGPT works Technically | ChatGPT Architecture [Video]. YouTube. https://www.youtube.com/watch?v=bSvTVREwSNw

EU AI Act: First regulation on artificial intelligence | News | European Parliament. (2023, 6 augustus). https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

George, A., George, A., & Martin, A. (2023). A review of ChatGPT AI's impact on several business sectors. Zenodo (CERN European Organization for Nuclear Research). https://doi.org/10.5281/zenodo.7644359

Hoes, E., Altay, S., & Bermeo, J. (2023). Leveraging ChatGPT for Efficient Fact-Checking. https://files.osf.io/v1/resources/qnjkf/providers/osfstorage/642b16668ef65b31d2429f1b?action=download&direct&version=2

Marchetti, A., Di Dio, C., Cangelosi, A., Manzi, F., & Massaro, D. (2023). Developing ChatGPT's theory of mind. Frontiers in Robotics and AI, 10. https://doi.org/10.3389/frobt.2023.1189525

Onslow, M. (2023, 5 mei). Understanding ChatGPT's parameters: Getting started. Medium. https://medium.com/@mike_onslow/understanding-chatgpts-parameters-getting-started-32ec12b5e51b

Padmavati. (2023, 17 augustus). ChatGPT in Business: Top 8 Transformative Uses. Northwest Executive Education. https://northwest.education/insights/careers/top-8-chatgpt-use-cases-for-businesses/

Russell, S., & Norvig, P. (2009). Artificial Intelligence: A Modern Approach (3rd ed.). Prentice Hall.

Ray, P. P. (2023). ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems, 3, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Shukur, B. S., Aljanabi, M., & Ali, A. H. (2023). ChatGPT: Exploring the Role of Cybersecurity in the Protection of Medical Information. Researchgate, 18–21. https://doi.org/10.58496/mjcs/2023/004

TheKnowledgeAcademy. (z.d.). Top characteristics of artificial intelligence You must know. https://www.theknowledgeacademy.com/blog/characteristics-of-artificial-intelligence/