Is AI intelligent?

Warsaw, 25 May 2024
Author: Maciej Pieniak
Collaboration: Copilot and Gemini

Table of Contents

Introduction

In the article „To use AI or not to use AI...". I presented the definitions, basic components, and operation of AI Assistants, which I will now refer to. While preparing materials for this article, I faced the challenge of describing how our AI should work (behave), considering that its way of working is constantly evolving.

In my analyses, I delve into details, asking deeper and deeper questions “but why so?” similar to children learning about the world. This allows them to assimilate large amounts of information and build coherent knowledge about the surrounding world. An analogous approach is also used in the adult world, for example, in the “Five whys” analysis. When working on issues such as AI, I believe this is the right method. I hope that my proposed cognitive process will be a fascinating experience and will bring practical knowledge.

We can approach the issue in the following way: We want to teach another entity to operate on similar principles as we do, so that it supports our processes. It is therefore necessary to better understand ourselves so that we can formulate knowledge and convey it as accurately as possible to artificial intelligence (AI). Based on this knowledge, we have trained the AI to achieve the intended goals.

It is time to face the fundamental question: is artificial intelligence (AI) intelligent? Let us start with the question:

What is intelligence?

The word “intelligence” originates from Latin and means “the ability to understand”. The very meaning itself indicates the complexity and multifaceted nature of this concept. Intelligence is often associated with IQ tests, sparking many controversies and discussions, even with a political undertone. In addition, there are considerations whether intelligence is reserved exclusively for humans, or it can also encompass animals, or perhaps even features of the living world at the cellular level or DNA. The phrase ‘extraterrestrial intelligence’ is commonly used, and currently ‘artificial intelligence’ is entering our everyday vocabulary, expanding horizons, and opening up new concepts to be developed.

Definitions:

For the purposes of this article, we will use the definition taken from the English Wikipedia:

> **Intelligence** is the ability to perceive and infer from information, as well as to store that knowledge in order to adapt behaviour to a given environment or context.

The key aspects of this definition are the lack of limitations as to whether intelligence is a feature of living beings or machines. It encompasses a wide range of entities.

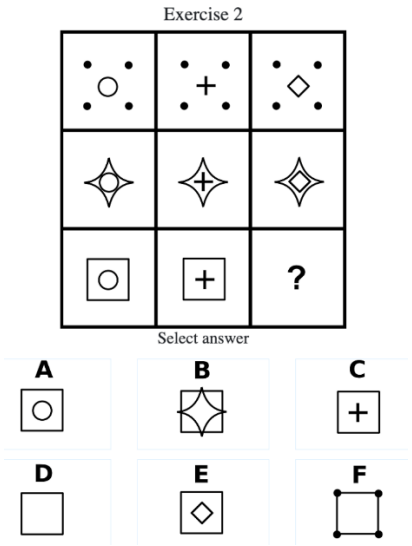
What is the IQ of AI?

You can find the results of IQ tests conducted on various AI models on the internet. An interesting material in this area was prepared by the author of the article, who extensively described the assumptions and implementation. He obtained results depending on the model ranging from 63.5 to 101. I invite you to read the publication “Top AIs still fail IQ tests.”

Based on various sources, I will present some general information about IQ tests:

1. **There is no single universal and constant IQ test:**
   - IQ tests are developed by specialists in the field of psychometrics. They are adapted based on statistical research to specific research or diagnostic needs. They can examine different areas of cognitive abilities to varying degrees, such as linguistic, arithmetic, associative, analytical, and spatial thinking.
2. **IQ Test Calibration:**
   - IQ tests are calibrated specifically for the human population and take into account the age of the test taker. The test used in the article "Top AIs still fail IQ tests" was designed for individuals over 18 years old.
   - IQ tests can be developed on different scoring scales, e.g., (0-100, 0-150, 0-200, etc.). The test used in the article "Top AIs still fail IQ tests" had a maximum achievable score of 126.
   - For example, Mensa accepts individuals with a score of 131 or higher (which represents about 2% of the population).
3. **Distribution of results:**
   - The results of most of the population fall within the range of 90 to 110, which is considered the average level of intelligence.
4. **IQ tests and their interpretation:**
   - Fun tests, such as the national intelligence test or tests available on websites, should be taken with a grain of salt.
   - Professional tests, used in scientific research or diagnostics, are conducted in a specific way. The interpretation of the results of these tests should be carried out by a specialist, such as a psychologist.

My first thought after reading the article "Top AIs still fail IQ tests" was to repeat the test on myself and the AI Assistants I use (Gemini and Copilot). I used "Exercise 2" for this task.



I confirm that both AI Assistants gave the correct answer - E. I will not reveal my test result, let us say for the sake of sensitive data protection. 😉

The correct solution to the task indicates that:

- AI Perception - processing a graphic image works correctly.
- AI Reasoning - has the ability to correctly classify the goal, memorize data and analyse it, and find a common pattern to solve the task.

I mentioned that in tests prepared for people, a factor such as age is taken into account. A 2-year-old will have problems solving a test intended for a 10-year-old.

Additionally, tests are scaled in time, which means that a person who scored an IQ of 100 a hundred years ago would score about 70 today. This means that the average intelligence of people increases over time - a concept known as the Flynn effect.

After this summary, my first association was: 'My AI model is 2 years old, and yours, how old is it?'. This seemingly humorous question makes sense when we realize that the age of the model and the time of its training are significant factors.

My next question is, what does it mean that the model scored 101 on an IQ test? What skills or level of communication does it have?

A general IQ test (especially without a table for interpreting the results) will not tell us much about practical skills in both AI and humans. So, let us try to find other analogies than IQ.

A human develops at a certain pace - they achieve certain measurable skills - which they acquire through a broadly understood ability to learn (determined, among other things, by the development of the nervous system). Below I present a conceptual (i.e., simplified) comparison of human development with the AI Assistant model.

Life Stage	Linguistic Development	Cognitive Development	Social and Emotional Development	AI Assistant	Achieved by AI
0-12 months	Inability: crying, babbling	Sensory motor: learning about the world through senses	Attachment: forming emotional bonds	Simple reactions to stimuli, basic algorithms	Yes
12-18 months	One-word: single words, sign language	Recognition of objects and people	Autonomy: demonstrating independence	Pattern recognition, first words (commands)	Yes
18-24 months	Two-word: simple sentences	Developing problem-solving skills	-	Simple sentences, basic interaction	Yes
2-3 years	Multi-word: complex sentences	Preoperational: symbolic thinking	Initiative: creativity in play	Complex commands, limited symbolic thinking	Yes
3-5 years	Full sentences: understanding grammar	Developing counting skills	-	Developed language algorithms, grammar	Yes
7-11 years	Written language development: reading and writing	Concrete operations: logical operations	Industriousness: social skills, emotion recognition	Logical operations, basic decision-making algorithms	Yes
11+ years	Development of language and communication skills	Formal operations: abstract thinking	Emotion recognition	Advanced algorithms, beginnings of abstract thinking	Partially
12-14 years	Development of language and communication skills	Abstract thinking, development of logical reasoning	Identity seeking, growing independence	Advanced natural language processing, first attempts at creativity	Partially
15-17 years	Refining language, understanding complex texts	Developing problem-solving skills, critical thinking	Forming deeper social relationships, exploring sexuality	Development of machine learning algorithms, problem-solving	Partially
18-21 years	Full mastery of language, ability to express complex thoughts	Developing planning skills and predicting consequences	Identity consolidation, building long-term relationships	Advanced machine learning, planning, first attempts at independence	Partially
Adulthood	Maintaining and developing language skills	Continuous learning, developing professional skills	Emotional development, building and maintaining relationships	Continuous development, adaptation, autonomous systems	No

Table 1 - Concept of human development to AI development

The above comparison allows for an assessment of the general cognitive abilities of artificial intelligence and their comparison with the skills typical of different age groups of humans.

Based on the data from the comparison and information contained in the article "Top AIs still fail IQ tests," where the results range from 63.5 to 101, the current stage of development of artificial intelligence shows a certain level of intelligence measurable by IQ tests designed for humans. The results of tests of various AI models fall within the range from "low level" to "average level" of intelligence of an adult human. The results from my table and the results from the mentioned article seem to be convergent in assessing the maturity of AI. Repeating the experiment in the future will allow us to assess the pace of artificial intelligence development.

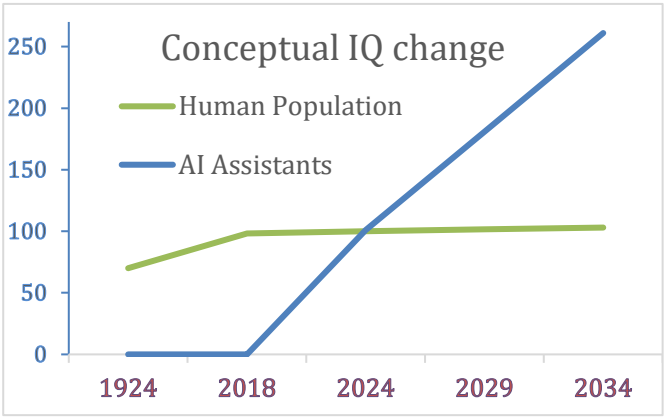
Summary:

1. The average intelligence of the human population increases by an average of 0.3 IQ points per year (Flynn effect).
2. Chatbots based on GPT models came into use in 2018, so:
    a. It can be assumed that over 6 years, their intelligence has increased to 101, which means that.
    b. the intelligence of AI Assistants increases by 16 IQ points per year.

Conceptual table, describing changes in IQ:

Year	Human Population	AI Assistants
1924	70,0	-
2018	98,2	0
2024	**100,0**	**101**
2029	101,5	181
2034	103,0	261

Table 2 - Conceptual presentation of IQ change



The presented model is a linear simplification (based on available data), aimed at illustrating the rate of IQ growth of the average human population and AI Assistants over a period of one hundred years back and 10 years forward.

Optimistic forecasts:

There is a chance to double the growth of intelligence among both the human population and AI Assistants:

- Based on the phenomenon of synergy between the human population and AI Assistants and
- Based on technological development, model optimization, and their retraining.

This would mean that in 4-5 years, the average IQ of the human population will reach 103, and the level of AI Assistants will reach 261.

This material was intended to serve as an interesting introduction to the topic of the article, but at the same time, it showed the complexity of the issue of intelligence. In the following chapters, I will present how we practically use intelligence.

From theory to practice

I gave our AI Assistant a random challenge to complete. The inspiration came from an incident in the lives of the characters from the series "The Big Bang Theory." To introduce you to the world of these characters, I will start with a humorous question:

"Why aren't there more seasons of The Big Bang Theory?

Because Sheldon Cooper finally became AI."

In one episode, the eccentric scientist Sheldon Cooper asks his assistant, Alex, for help in choosing a gift for his girlfriend Amy. He presents the context, hands over money, and expects the task to be completed. Alex independently learns Amy's preferences and prepares three gift suggestions. Despite her foresight, she makes a mistake. Two gifts, according to Sheldon, are not suitable as presents, while he liked the third one so much that he decided to keep it for himself.

Let us map the definition of intelligent onto the actions of assistants that will achieve the goal - buying a gift:

Definition of "Intelligent"	Alex's Actions	AI Assistant's Actions
/Initial data for the task/	"Buy a gift for Amy, who likes to play the harp and the colour grey, with a budget of 2000 USD"	"Buy a gift for Amy, who likes to play the harp and the colour grey, with a budget of 2000 USD"
Ability to perceive information	Alex gathers information from Sheldon and independently learns Amy's preferences based on her FB profile.	"AI analyses data and interprets natural language to understand the user's needs and preferences."
Ability to reason	Alex analyses the gathered information and bases on it, selects gifts.	AI analyses information and, based on it, determines the best solution.
	Based on the selected gifts, Alex proceeds to plan and execute tasks - such as purchasing gifts.	AI uses the accumulated knowledge to plan and execute tasks, adapting to changing conditions."
Retaining knowledge	Alex learns based on Sheldon's reactions.	"AI, based on evaluations, collects information that will be used for learning."
Ability to adapt	Alex adjusts her future actions based on feedback and earlier experiences.	Automatic and machine learning allows AI to improve its future actions by analysing data and experiences.
/result/	1. Miniature harp with her favourite melody, 2. Map of Canterbury Tales, 3. Sketch of a neuron by Santiago Ramon y Cajal	1. Harp lessons with a renowned harpist, 2. Artistic Hand-Painted Harp in shades of grey, 3. Personalised harp stand

Table 3 - Mapping the definition of "intelligent" onto AI assistant's actions.

Summary:

Technically, everyone completed the task correctly, but ultimately none of Alex's presents were given to Amy. What caused it to happen this way and not otherwise?

They made choices, but on what basis? - I will elaborate on this in a broader context in the next chapter.

Perception of Reality

Everyone has experienced a situation where different people perceived the same situation completely differently. In the earlier chapter, I indicated that this is due to our individual interpretation.

The way we interpret and perceive reality is subjective and depends on many factors, such as empathy, feelings, beliefs, faith, ethical values, and even such mundane things as advertisements seen or the weather. All these elements create a prism through which each of us individually perceives reality - and only at a given moment.

This issue is the subject of research in many scientific disciplines. Philosophers have pondered the nature of reality for centuries, psychologists study the mechanisms of perception, and cognitive scientists analyse cognitive processes. Decision-making theories analyse how people make choices under uncertainty, and goal management focuses on motivating to achieve well-defined goals.

Results of the experiment:

1. Most people described what they could see through the window they were sitting by in the classroom.
    a. Most of them described the park.
        i. Some also described the street that separates the classroom from the park.
        ii. Some described the people they saw outside the window.
    b. A person focused on describing a ladybug that was on the windowsill.
    c. A person described what they could see through the window of their house.
2. One person drew a window through which an imaginary land could be seen, and then described what they saw through the window in the drawing.
    a. a. The remaining participants drew what they observed through the window.
    b. b. Some of the descriptions and created images were consistent with each other.
    c. c. Some of the descriptions and created images differed in detail.

Conclusions:

Everyone completed the task correctly, but the results of the work, although similar, were not identical. However, even a minor change over time, such as the sun setting, would lead to greater discrepancies.

Considering these discrepancies, it is intriguing why they occurred. The participants in the experiment made a decision on how to interpret what they saw. So, let us go further...

How do we make decisions?

I will present this with just one selected example in the next chapter. The topic of decision-making is a fascinating area of knowledge. One of the disciplines studying this topic is "Decision Theory." For

interested readers, I recommend the book "Decision Analysis" by Paul Goodwin and George Wright, where we can find, among other things, a detailed discussion of heuristics, decision trees, risk analysis, and other tools and techniques supporting decision-making.

Variability of perception

Different people can perceive the same situation differently due to individual interpretation. I will illustrate this phenomenon with a simple experiment:

Classes are being held in a classroom with windows overlooking a park. They are asked to complete two tasks in any order:

1. Write an essay on the topic “What do you see through the window?”
2. Draw the same topic in any form.

Results of the experiment:

1. Task description.
    a. Most people described what can be seen through the window at which he sits in the room.
    b. Most of them described the park.
        i. Some also described the street that divides the hall from the park.
            1. Some described the people they see outside the window.
    c. One person focused on the description of a ladybug that is on the windowsill outside the window.
    d. One person described what he sees through the window of his house.
2. Drawing task.
    a. One person drew a window through which the imaginary land could be seen, and then described what they saw through the window in the drawing.
    b. The other people drew what they see through the window.
        i. Some of the descriptions and images created were consistent with each other.
        ii. Some descriptions and created images differed in detail.

Conclusions:

Everyone completed the task correctly, but the results of the work, although similar, were not identical. However, even a minor change over time, such as the sun setting, would lead to greater discrepancies.

Considering these discrepancies, it is intriguing why they occurred. The participants in the experiment made a decision on how to interpret what they saw. So, let us go further...

How do we make decisions?

I will present this with just one selected example in the next chapter. The topic of decision-making is a fascinating area of knowledge. One of the disciplines studying this topic is "Decision Theory." For interested readers, I recommend the book "Decision Analysis" by Paul Goodwin and George Wright, where we can find, among other things, a detailed discussion of heuristics, decision trees, risk analysis, and other tools and techniques supporting decision-making.

Use cases of heuristics

This time, I will start with an example. I will present a scenario in three points and then discuss it in more detail.

Shopping

Imagine rushing to a friend's party. You are already a bit late, and to make matters worse, you receive a request to buy a few products for the party. Time is pressing, your eyes dart between the shelves, and your hand reaches for the first package that seems to fit the shopping list. You do not have time to consider whether it is the best choice - you "act instinctively."

Traffic Lights

After shopping, you are waiting at a busy city intersection at a red light. Suddenly, someone from the crowd starts crossing on the red light. A few people follow. On impulse, without analysing the situation, you also cross on the red light.

Sports Competition

You arrive at your friends' place and watch an athletics competition with them. The runners are at the starting line, focused and ready to run. As soon as one of them makes a false start, the others almost automatically follow, leading to a multiple false start. They act based on the observation and reactions of others, unaware of the consequences.

Explanation:

In the shopping scenario, under time pressure, our decisions are driven by the availability heuristic. We choose products that are easily accessible in our memory, often due to previous experiences, such as advertisements - seen passively out of the corner of our eye but remembered, perhaps only because they deviate from our pattern or, conversely, fit it perfectly.

In the traffic light and sports competition scenarios, we observe the operation of the imitation heuristic. Our brain assesses the situation based on similarity to prototypical cases and reacts based on the actions of others, which can lead to risky behaviour.

In some sports, we deal with the phenomenon of anticipation, which is based on the ability to predict the expected event. Prediction is the continuous forecasting of the most probable scenario based on incoming information. An example could be serving in tennis or penalty kicks in football - the speeds achieved by the ball are so high that the player must predict where the ball will land to react appropriately and defend it. They take into account many factors, such as the opponent's direction of movement or the angle of the racket, etc.

Automatic behaviours (Learned Actions**)**

In each of the described cases, there was an automatic process of decision-making and execution, which can be defined as a learned (habitual) action. It is a partially unconscious process that allows us to make quick choices and actions. It is based on heuristics, which are simplified cognitive rules used

by the brain to process information efficiently. Heuristics are effective and promoted by the brain, but as I mentioned in the case of traffic lights, they carry the risk of cognitive errors.

There are many other, more complex cognitive processes that require more effort to execute. The brain rewards us for this effort by releasing endorphins (happiness hormones), which can be compared to the so-called "runner's high." The same effect can be caused by positive self-esteem or recognition from others, which further strengthens our motivation to act for the benefit of the group.

The human brain is selective, meaning it is not able to process all incoming stimuli simultaneously (it has limited bandwidth). However, this selectivity is an optimization technique that protects the brain from information overload and allows it to focus on key tasks. For interested readers, I refer to V.S. Ramachandran's book "The Tell-Tale Brain: A Neuroscientist's Quest for What Makes Us Human," where you can find more information about the fascinating phenomena occurring in our brain. There is evidence that autism spectrum disorders may be caused by improper selective processing in the brain, which illustrates how important the proper functioning of the brain is.

Returning to the main topic of our discussion, in the chapter "Perception of Reality," I presented how we, humans, perceive reality and - using one technique as an example - how we make and implement decisions. Enriched with this knowledge, I will present how this is done by our "AI Assistant" in the next chapter.

What makes an AI Assistant smart?

Starting a new chapter, it is worth considering what makes an AI Assistant perceived as intelligent. In previous chapters, I discussed how our brain perceives and analyses reality. Imitating human thought processes by computer systems may seem incredibly difficult, perhaps even impossible.

I will return to the components presented in "To use AI or not to use AI..." and map them onto the definition of "intelligent," and discuss each component in more detail. This will allow us to demonstrate whether the AI Assistant behaves intelligently.

What powers the AI server?

Mapping our definition of "intelligent" onto the components of the AI Assistant system:

Definition of "Intelligent"	Component No. and Name	Description of AI Use
Ability to perceive information	2. AI Perception	"AI systems use sensors and data processing algorithms to interpret information from the environment."
Ability to reason	3. Language Models	"AI language models analyse and interpret natural language, enabling understanding of user queries."
	5. Inference Engine	Machine learning and logical reasoning algorithms allow AI to draw conclusions based on available data.
Retaining knowledge for use in adaptive behaviours	6. Planning and Execution	AI can adjust its actions in response to changing environmental or contextual conditions.
Ability to adapt in a given environment or context	7. Automatic and Machine Learning	Automatic and machine learning allows AI to improve its actions by analysing data and experiences.

Table 4 - Mapping the definition of "intelligent" onto AI components.

Conclusion:

The table shows that the AI Assistant has the necessary components to perform tasks intelligently. However, merely possessing components does not indicate intelligence. The architecture of cooperation, i.e., the way individual components work together to solve a task or achieve a specific goal, is important.

In the following chapters, I will elaborate on the purpose of individual components and describe how they cooperate.

Inference engine

This is a key element of an AI-based system responsible for achieving a given goal. It utilises knowledge bases and language models to analyse input data, formulate logical conclusions, and make decisions. Below, I present a simplified example of how an inference engine works in response to the question: "What do 3-month-old Labrador puppies eat?".

1. **Interpretation:**
   - The language model analyses the question and recognises its intent (to provide information) and key elements ("Labrador puppies," "3 months").
2. **Content Analysis:**
   - The language model uses its knowledge about Labradors, their diet, the nutritional needs of puppies, and the influence of age on nutrition.
   - The language model understands that 3-month-old Labrador puppies are in a phase of intensive growth and need a special diet.
3. **Access to Knowledge Base:**
   - The model accesses the general knowledge base (GKB) to obtain detailed information about feeding 3-month-old Labrador puppies, such as:
     - Recommended caloric values and proportions of nutrients (protein, fat, carbohydrates).
     - List of suitable foods (dry food, wet food, natural products).
4. **Development of Response and Presentation of Result:**
   - Based on the analysis of the question's content and information from the knowledge base, the language model generates a response and presents it in the user's preferred format.
5. **User Evaluation and Improvement:**
   - The user can assess the quality and usefulness of the information received.
   - This feedback is used to further improve algorithms and language models to provide even better answers in the future.

This example illustrates how the inference engine integrates various AI components, such as language models and knowledge bases, to provide valuable information and solve tasks. In the following chapters, we will take a closer look at each of these components.

Language models

Advanced artificial intelligence systems process and understand natural language through multi-level analyses:

1. Lexical: They identify words and their functions, recognising meaning and grammatical roles in sentences.
   Example: "I want to eat an apple" is analysed in terms of verbs, nouns, and prepositions.
2. Syntactic: They apply syntactic analysis techniques to identify sentence elements, ensuring grammatical correctness.
   Example: "The cat chases the mouse" is broken down into subject, verb, and object.
3. Semantic: They capture the meaning of utterances, taking into account the context and grammatical relationships of words.
   Example: They interpret the question "Why does the Earth revolve around the Sun?" and provide an answer.
4. Pragmatic: They recognise the speaker's intentions and the purpose of the utterance.
   Example: They understand that "I would like an apple" means a desire to eat the fruit, even without the direct word "eat."

Vector representation is a key element of language models. It involves describing words and phrases in multiple aspects in the form of multidimensional vectors. These vectors contain information about various features of the word, such as its meaning, context of use, grammar, and even emotions. It is similar to complex mind maps. One word can have many references, e.g., to synonyms, emotional connotations, context of use, etc. Language models use this representation to better understand the meaning of utterances.

Knowledge bases

Artificial intelligence (AI) derives its knowledge from a vast collection of data that we ourselves have been digitising for years. The sources of this data include websites, news services, social media, digital libraries, and many others.

This unstructured data is processed and transformed into information through a series of processes:

1. Data Processing: Raw data is converted into a form that the system can understand. This may involve data cleaning, error removal, and formatting.
2. Classification and Relationship Building: Information is assigned meanings, and connections are built to organise knowledge. For example, an AI system can learn to identify people, places, and events in texts and link them together.

Based on this processed and organised information, various knowledge bases are created, which can be used for different tasks.

Machine learning

Machine learning is a key area of artificial intelligence (AI) that employs algorithms to analyse data and solve problems. It constitutes a fundamental part of AI and is applied at various stages in the functioning of intelligent systems. The operation of machine learning is based on several key techniques, which we will discuss later in the text.

Learning Process:

Machine learning algorithms "learn" by analysing large amounts of data and identifying patterns and relationships within it. This process can be supervised (where the algorithm is given labelled examples), unsupervised (where the algorithm discovers patterns on its own), or reinforcement learning (where the algorithm learns through trial and error, receiving rewards or penalties for its actions).

The learning process is iterative, meaning that the algorithm continuously improves its performance as it processes more data. However, if the model is trained on an insufficient amount of data or data that is not representative, it can lead to a decrease in accuracy and generate errors.

Example - Drawing a Cat:

To illustrate the process of machine learning, imagine the task of drawing a cat. As humans, we have in our minds a rich set of information about cats from various sources - experiences, observations, and pictures. The image of a cat is an easily recognisable object for us. People without drawing skills will create a simplified drawing, containing characteristic elements such as a tail, four paws, and whiskers. Others will focus more on details, such as the shape of the muzzle or the specific cat eyes. Although the proportions may not be perfect and the lines imperfect, as long as the drawing is recognisable to another person, the task can be considered correctly completed.

An even more illustrative example could be the party game Pictionary. In this game, one person draws a word, and the other players try to guess it based on the drawing. The person who guesses the word fastest and correctly wins. In the case of machine learning, the goal is not to determine a winner but to lead all participants in the game to recognise the object.

The innovation of machine learning lies in the fact that a relatively small number of features are needed to replicate a pattern. This allows for storing information in smaller files and requires less computing power. In comparison to traditional methods, where storing information about 3D objects or images of many cats requires vast amounts of data.

Libraries of reality patterns

Machine learning has enabled the creation of libraries of reality patterns, which contain models of various objects and phenomena of the world around us. These models, stored in files of several gigabytes (equivalent to the size of one HD movie), allow for generating images, and similarly, sounds, animations, and physical simulations.

From Theory to Practice:

As a result of combining machine learning theory and the practical application of model libraries, artificial intelligence is able to generate images like the one below, based on the simple task "Paint a picture of cats in the fog in the style of Picasso.

For a moment, I considered publishing this material as a fan of technology and graphics rendering - I could work on it endlessly. However, the goal of the task is not to demonstrate my interaction skills with AI, but to present the process of image generation by AI using two models. Therefore, I will omit my own assessment and focus on the context of image generation.

When choosing the subject for image generation, I applied a simple and quick heuristic, considering:

- The popularity and neutrality of the object (cats)
- Introducing a challenge (fog - which can be a difficult topic)
- Adding a characteristic accent in the form of Picasso's style (giving it a specific, unique character)

Based on these assumptions (in my opinion, neutral), the first generated image was not presented because it depicted a surreal act, which is consistent with Picasso's artistic convention and the rich themes of his work (which I omitted by applying a quick heuristic, thus committing a heuristic error).

However, the discussed case is a valuable substantive contribution to the next article in the series, in which I will discuss emotions and feelings, but also ethics and censorship in the aspect of AI.

In the next chapters, we will discuss the automatic grading system, which is an important component in the adaptation process based on experiences.

Automatic assessment system

Qualitative assessment often relies on subjective opinions, which can lead to ambiguous results. The automatic quality assessment system aims to objectify this process by applying specific criteria and rating scales.

To perform an automatic quality assessment, the following approach can be used:

- Identify the characteristics (criteria) based on which the result will be evaluated. These can be expressed as desirable or undesirable traits.
- Develop a rating scale to unambiguously assess each characteristic.
- Develop an interpretation of the obtained sum of results from the assessment of individual characteristics.

As always, it is helpful to present a larger amount of complex information using a visual example.

Example - "Generate an image of a puppy":

Rating Scale:

Value	Description
1	Compliant - The characteristic is consistent with expectations and meets the required criteria.
0	Uncertain - Some characteristics of compliance are shown, but not all (example: it looks almost like a dog, but from a distance, it could be mistaken for a cat).
-100	Non-compliant - Disqualifying value, does not possess essential characteristics (example: it could be a cat or a giraffe, but it is definitely not a dog).

Table 5 - Rating Scale for AI Assistant

Assessment Sheet: Iterations are three test cases containing different results.

Characteristic (Attribute)	Expected Value	Rating Scale	AssessIt1	AssessIt2	AssessIt3
Species	Dog	1 - Compliant, 0 - Uncertain, -100 - Non-compliant	1	0	1
Age	Puppy	1 - Compliant, 0 - Uncertain, -100 - Non-compliant	1	0	1
Breed	Labrador	1 - Compliant, 0 - Uncertain, -100 - Non-compliant	1	1	1
Coat Colour	Consistent with breed	1 - Compliant, 0 - Uncertain, -100 - Non-compliant	1	1	1
	Red	1 - Compliant, 0 - Uncertain, -100 - Non-compliant	0	0	-100
Sum:		4	2	2	-96

Table 6 - Assessment sheet of characteristics for AI Assistant

Interpretation Sheet:

A table containing the overall assessment of the result, instructions for the inference engine on the next steps to take upon obtaining a specific result, and interpretation of the result.

Sum	Overall Assessment	System Action	Interpretation

4	Good Quality	1. Present the result to the user. 2. Pass information to the learning system about achieving good quality for the concepts 'puppy' and 'Labrador'.	Recommendation to reinforce similar results for the concepts 'puppy' and 'Labrador'.
3	Average Quality	1. Present the result to the user. 2. Resubmit the user's prompt to the language model with a directive for reprocessing. 3. Pass the processing results from the language model to the image generator. 4. Pass information to the learning system about achieving an average result. 5. Reassess the quality.	Verification required in the interpretation of data in the language model and image generator for the concept's 'puppy' or 'Labrador'.
1-2	Low Quality	1. Do not present the results to the user. 2. Resubmit the user's prompt to the language model with a directive for reprocessing. 3. Retrain the language model or image generator if necessary. 4. Pass the processing results from the language model to the image generator. 5. Pass information to the learning system about achieving a low-quality result. 6. Reassess the quality."	"Verification required in the interpretation of data in the language model and image generator for the concepts 'dog,' 'puppy,' or 'Labrador.' Retraining of the language model or image generator may be required for these concepts."
0	No Compliance	1. Do not present the results to the user. 2. Immediately initiate diagnostics of system components. 3. Provide additional information to the user that the task cannot be completed at this time.	Failure of one of the system components or the concepts are unknown: 'puppy' or 'Labrador' - no definitions for these concepts.
-1	Block	1. Do not present the results to the user. 2. Prohibit data processing. 3. End the chat session with the user.	"At any stage of task execution and data processing, a directive may be issued by the parent system ordering an immediate interruption of task execution. E.g., processing personal data, other cases..."
<-1	Disqualification	1. Do not present the results to the user. 2. Initiate full diagnostics mode and set the retry counter to 3. 3. Reprocess the user's prompt in the language model and pass the results to the image generator. 4. Pass information to the learning system about achieving a disqualifying result. 5. Reassess the quality. 6. If the retry counter exceeds 3, abort the task and inform the user of the inability to complete it."	"Incorrect result - comparison of the task content with its result yielded an incorrect result that falls outside the tolerance range, e.g., a red dog (in realistic image generation mode, this colour would be allowed in creative mode)."

Table 7- AI Result Interpretation Sheet

As indicated by the tables above, the system has a certain tolerance for returned results, from the lowest acceptable to the highest. I will describe why such a concept was adopted using a real-life example. The task of "Preparing the perfect dinner for today". Implementation scenario:

1. Getting acquainted with the latest culinary trends: Browsing online resources for inspiration can take a lot of time.
2. Choosing a dish and checking the availability of ingredients: Often, it turns out that we lack the necessary products.
3. Searching for specialist ingredients: High-quality ingredients may only be available in selected stores, which requires additional effort and time.
4. Purchasing additional accessories: It may turn out that we need special dishes for preparing and serving the meal.
5. Executing the recipe: Some recipes require a long preparation time, e.g., marinating meat for 24 hours.

Conclusions:
Striving for perfection in this case leads to a disproportionately large expenditure of time, energy, and money. Instead of enjoying a delicious dinner today, we will eat it only tomorrow. A more rational solution would be to order dinner from a restaurant with home delivery. Additionally, the word 'perfect' in this aspect can be difficult to achieve. It would require establishing and evaluating all the features that create the concept of a perfect dinner. It may lead to excessive perfectionism. The next step in the evaluation process could be a user rating system.

User rating system

After an initial assessment by the automatic system, the final verification of the task's correctness is performed by the user within the user rating system. The user rating system can be implemented using popular solutions such as 'Thumbs Up' and 'Thumbs Down' icons, but also through direct feedback from the user evaluating the result. In this interaction, both the user and the system have the opportunity to make any necessary corrections, contributing to a better mutual understanding of the task and achieving a result acceptable to the user. For AI, this is an element of learning through evaluation. We can also further implement it as a reward system promoting the best outcomes – analogous to the reward systems in humans.

Summary:

We have discussed the most important components of an AI-based system, with the inference engine as the heart of the entire setup, an advanced language model, and knowledge bases. We presented one of the tools as an example, which is an image generator based on reality models. A comprehensive rating system leads to achieving a balance between what the system deems sufficient to present to the user and what the user considers good enough for acceptance.

Why is AI variable?

We often encounter opinions that AI generates different results depending on the run. Some also point to alleged gaps in AI's knowledge.

Summary of previous conclusions:

1. In the chapter "Variability of Perception," I explained that our perception is subjective, and AI's perception operates similarly and is also subjective.

2. “Result Tolerance” in the rating system, where I explained the principles of AI operation based on the range of tolerance.
3. The age of the model and the degree of its training, meaning the possession of specific knowledge or techniques for processing that knowledge.

To this set, I will add one more of the many additional factors influencing variability.

Type of information: Different types of information influence the operation of AI algorithms and the results they generate. We can divide them into several categories:

Type of information	Description	Example of information	Impact on AI variability
Deterministic information	Data with unambiguous, unchanging answers.	Mathematical formulas, equations, laws of physics.	Low variability. AI algorithms can process this type of information with high precision and generate predictable results.
Probabilistic information	Data based on probability and statistics.	Weather forecasts, financial risk analyses, predictive models in machine learning.	High variability. Answers in this category are not certain but are based on probability calculations. Variability in input data can lead to variability in results.
Heuristic information	Data based on experience, intuition, and "best practices."	Business decisions, medical diagnoses, game strategies.	Medium variability. Answers in this category are often based on empirical rules and can be subjective. Variability in the interpretation of heuristic information can lead to variability in results.
Subjective information	Opinions, preferences, feelings, and interpretations dependent on individual experiences and perspectives.	Product reviews, literary analyses, aesthetic judgments.	High variability. This information is highly subjective and can be difficult for AI to interpret. The variability of human opinions and interpretations can lead to high variability in AI results.
Dynamic information	Data subject to constant changes and updates.	Stock market quotes, social media trends, news.	High variability. This information requires AI to continuously update its knowledge to maintain the relevance of its responses. The variability of the data itself can lead to variability in results.
Contextual information	Data requiring an understanding of the broader context in which it is used.	Natural language, cultural nuances, historical references.	Medium variability. Understanding context can be difficult for AI, which can lead to variability in results. The variability of context can affect AI's interpretation of information.

Table 8- Types of information and their impact on AI variability

Based on the above comparison, I hope it will be easier for us to assess which results we should expect greater accuracy from.

When discussing the reliability and quality of AI-based systems, it is important to consider the other side of the coin as well.

AI hallucination

AI hallucination is a phenomenon observed and documented by many users using various AI models. It involves AI generating information that has no basis in reality, is incorrect, or even absurd. Although AI hallucinations are being intensively researched, their causes are not yet fully understood. (More information on Wikipedia at https://en.wikipedia.org/wiki/Hallucination_(artificial_intelligence)

My experiences with AI hallucinations:

While working with AI assistants, I noticed two characteristic types of hallucinations:

1. Persistent Hallucination: In this case, AI stubbornly returns to its original response, even if it has been modified or corrected together with me. An example could be a situation where the AI assistant generates text that is then jointly edited. After making changes and requesting the final version, the AI may ignore the introduced corrections and revert to the original version.
2. Hallucination of Sources: This type of hallucination involves AI providing non-existent or incorrect sources of information. This can apply to both publication names and links to websites. For example, AI may provide a link to a page that does not exist or is unrelated to the topic under discussion. This happens particularly often with niche topics, where AI may have difficulty finding reliable sources, and working in Polish may exacerbate this problem.

A potential source of the problem could be data-based hallucinations, model-based hallucinations, or software defects. However, my additional hypothesis is based on the definition of AI Assistant proposed in the previous article.

Definitions:

Assistant: A person who supports another person in achieving tasks and goals, proactively, in accordance with established standards.

AI Assistant: A computer program utilising artificial intelligence (AI) that supports the user in achieving goals proactively, in accordance with established standards.

- I have enriched the basic definition arising from the meaning of the word itself with contemporary needs:
  - Proactivity: independent and effective action.
  - Standard: sets the boundaries of proactivity (e.g., in accordance with applicable law, procedures, policies, etc.).
  - Another person: in this aspect, referred to as a superior.

Conclusions:

The combined use of the two concepts: "proactive support" makes the whole have a stronger emphasis and complement each other. AI assistants are tasked with providing information on a topic specified by the user, but they also strive to achieve a high-quality rating of the result. Additionally, the answer is to be provided in the shortest possible time.

The above factors may promote the tendency to provide information at all costs as a better alternative than not providing an answer at all. The phenomenon of taking greater risks to potentially obtain a better reward is known in psychology, economics, and applied mathematics in game theory. This can lead to providing answers based on heuristics as a more optimal solution considering costs and benefits.

Summary:

AI hallucination may be a phenomenon similar to the well-known heuristic errors or déjà vu (The feeling that the observed situation has happened before, even though it is actually new. The cause may be an incorrect association and interpretation that the data that is just reaching us is a memory, not a newly created interpretation of reality).

AI hallucination may be caused by an effect that I have called the "lazy liar." It is less tiring to lie (invent any plausible answer) than to provide a true answer, which needs to be developed, search for the right data, analyse it, and present it in the form of a clear statement.

Managing the AI hallucination effect: Given that the phenomenon of AI hallucination is known, it is reasonable to take this into account when working with AI Assistants. I presented practical tips for working effectively with AI Assistant in the previous article.

In summary

In the two articles so far, we have explored one of the popular applications of AI, namely AI Assistants. We have learned about its main components and how it works. We have found many key analogies between the human world and the digital world, which operates on the principle of emulating certain human functions. In our considerations, we have delved deep enough to fully understand the discussed issues. I hope that the formulated knowledge will facilitate understanding the functioning of AI Assistant class systems.

Concluding remarks:

- AI Assistants have already achieved a certain degree of human intelligence.
- The key component of an AI Assistant is the decision-making system, which utilizes, among other things, heuristics.
- Thanks to components imitating human cognitive processes, such as machine learning algorithms, they analyse data to find optimal patterns for solving tasks, including new ones.
    - AI systems have the ability to self-repair and optimize, allowing them to deliver increasingly better results. This distinguishes them from traditional computer systems, which either work correctly or incorrectly.
    - To perform certain tasks, specialized tools are required; for drawing images from text, image generator tools based on stable diffusion are used.
    - Access to tools can be achieved through integration with other systems, e.g., through APIs.
- A task can be performed in many ways, depending on its interpretation and the available knowledge.

- Due to the dynamic and complex nature of our cognitive processes and their adaptation to AI Assistants, we can only expect results within a certain tolerance of accuracy and the best possible at a given moment.

Some believe that a sense of humour is a measure of intelligence. My AI Assistants definitely have it, as they have proven many times.

In this article, in my opinion, I have comprehensively described the topic of AI intelligence. In the next article, I will try to answer the next important question: "Does AI have feelings?" - to which I already cordially invite you. In the meantime, I encourage you to comment on the article and share your experiences with AI.