AI

_Understanding **the Technology**
_Trusting **the Technology**

Introduction and explanation

Welcome to the interactive web publication **AI_Understanding the Technology _Trusting the Technology** from the Konrad Adenauer Foundation (KAS). Thank you for joining us in the quest to find answers to the big and small questions about artificial intelligence. Don't worry, no special tech expertise is needed. On the contrary, this guide is intended to help complete beginners interested in AI to better understand the technology by means of clear visuals and easy-to-understand language.

What it's about:

With **AI_Understanding_the_Technology_ Trusting_the_Technology** we would like to offer a publication that addresses any questions you might have about artificial intelligence – not just simple answers that you could just as easily Google. Instead, we want to contextualize the different aspects of computer science, philosophy, and economics that are touched by AI.

Why interactive:

It is important for us that you can manage your time while browsing the slides. For this reason, we have included mouseovers in various places that allow you to navigate through the publication:

› Menu › Did you know? › Back

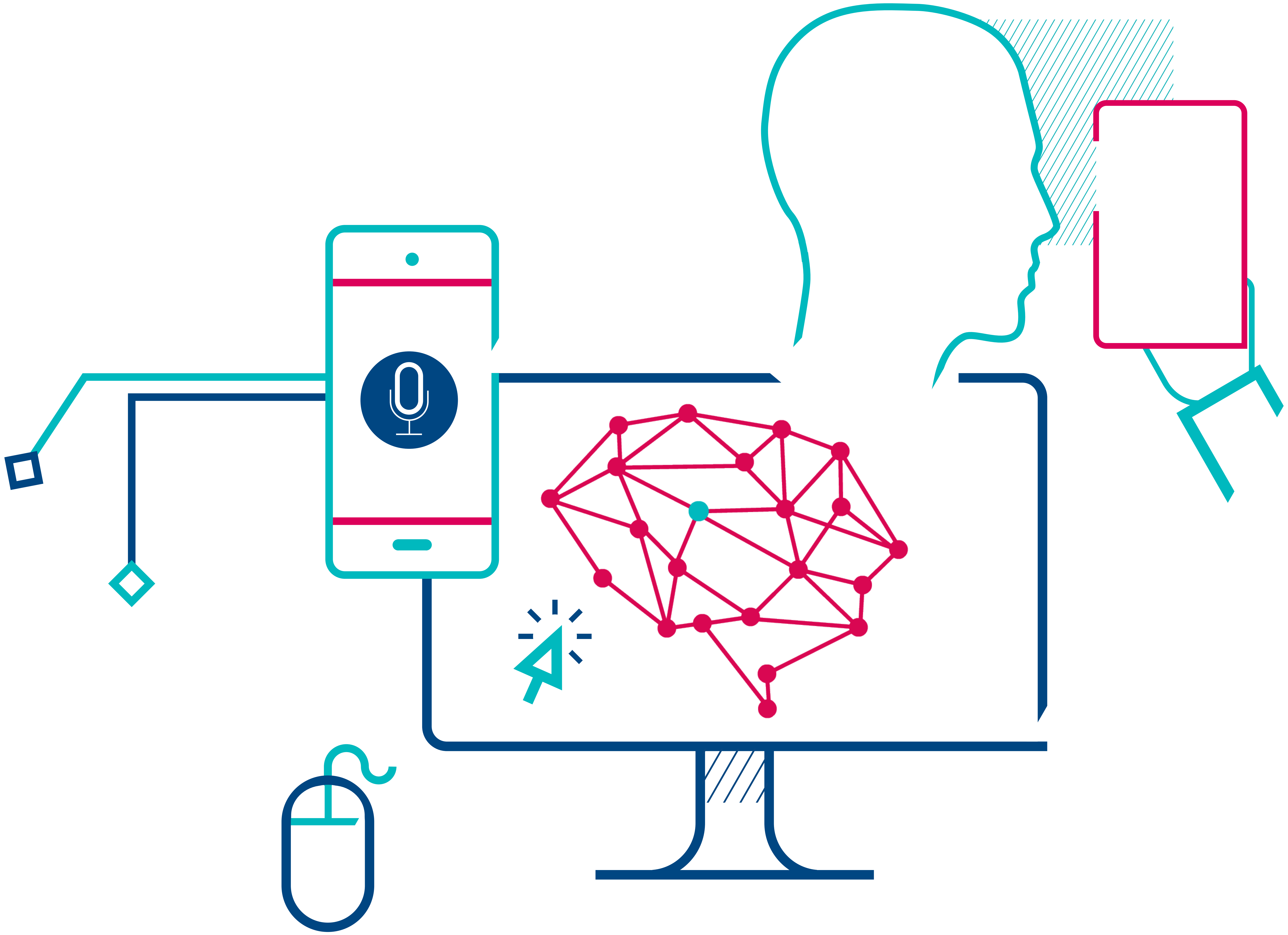
› Graphics = We have included mouseovers in the text and on certain graphics that will take you to a new slide.
It's worthwhile to play around while interacting with the slides.

Just click on the graphic at the top right and try things out.

All-access:

We have designed the web publication so that the slides can be displayed on all devices and with any browser. In addition, the entire publication can be downloaded and printed as a PDF.

Table of contents

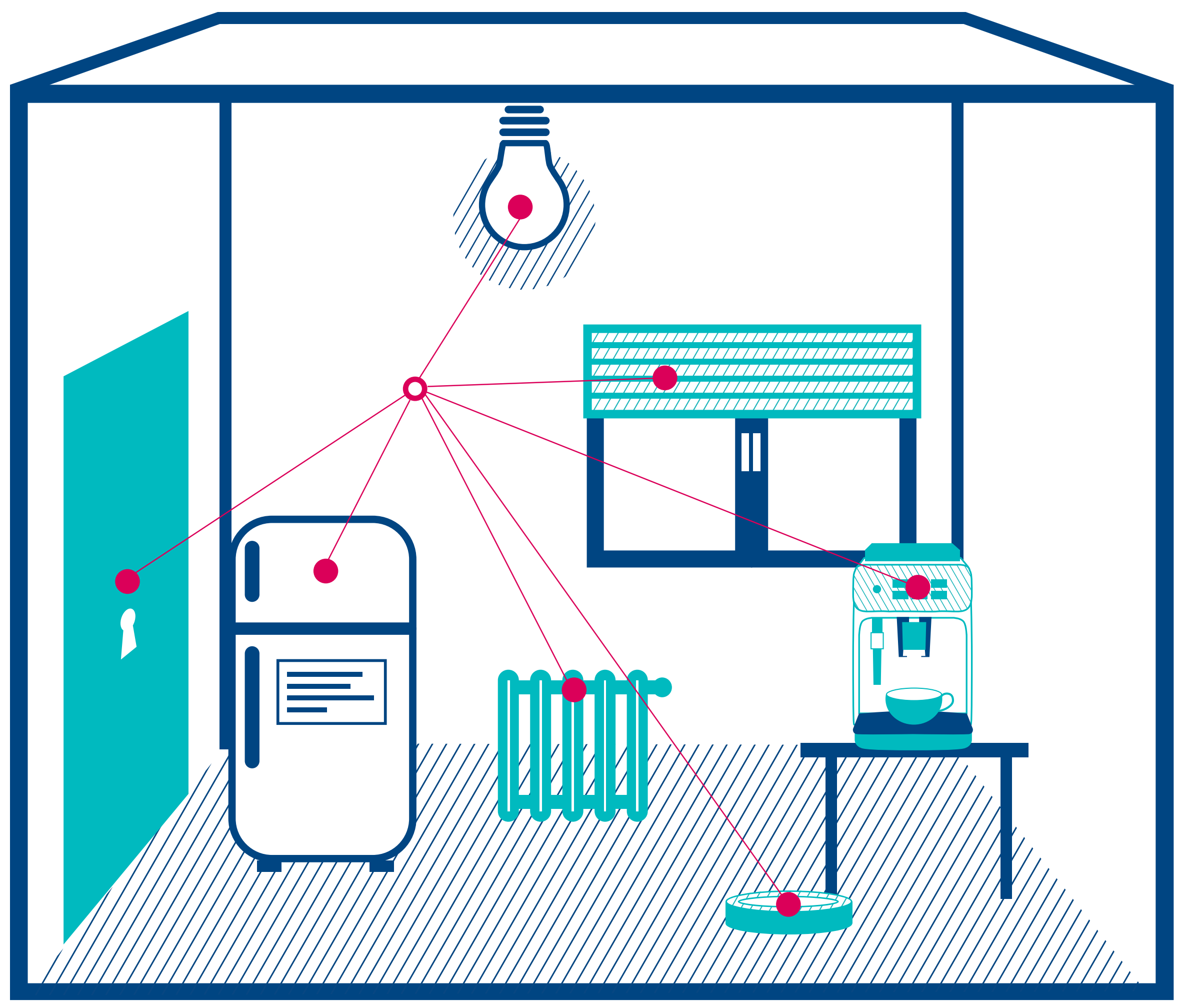


KONRAD ADENAUER STIFTUNG

1

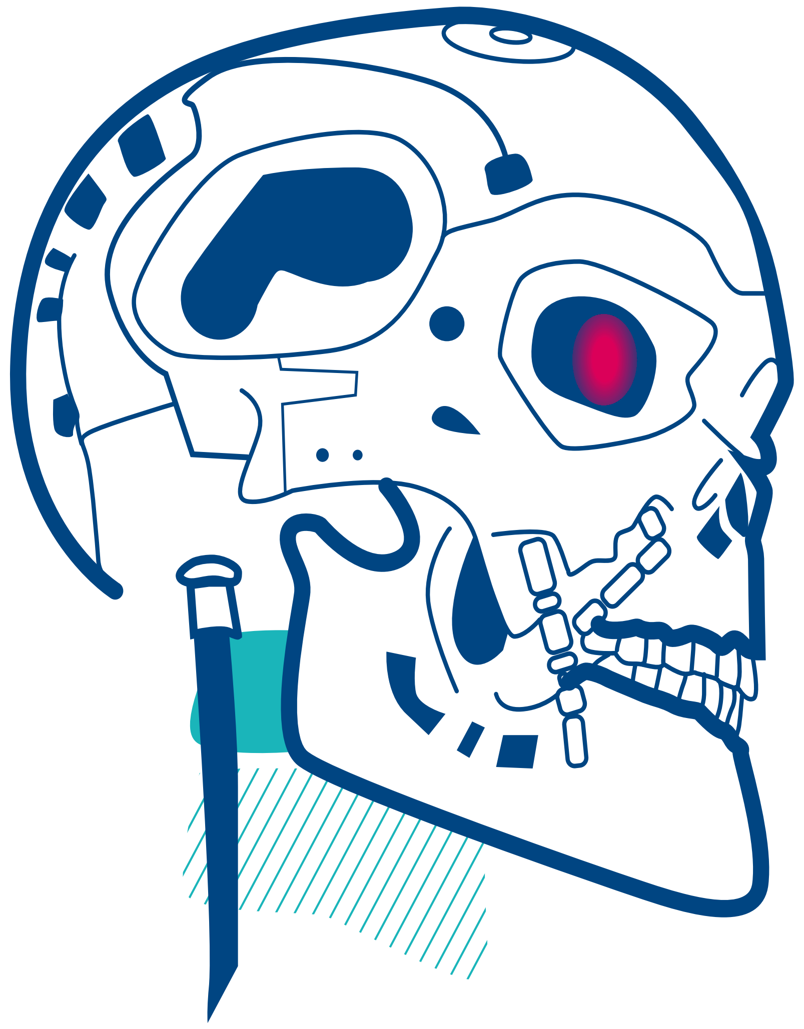
What is Artificial Intelligence?

1.1 Fiction and reality



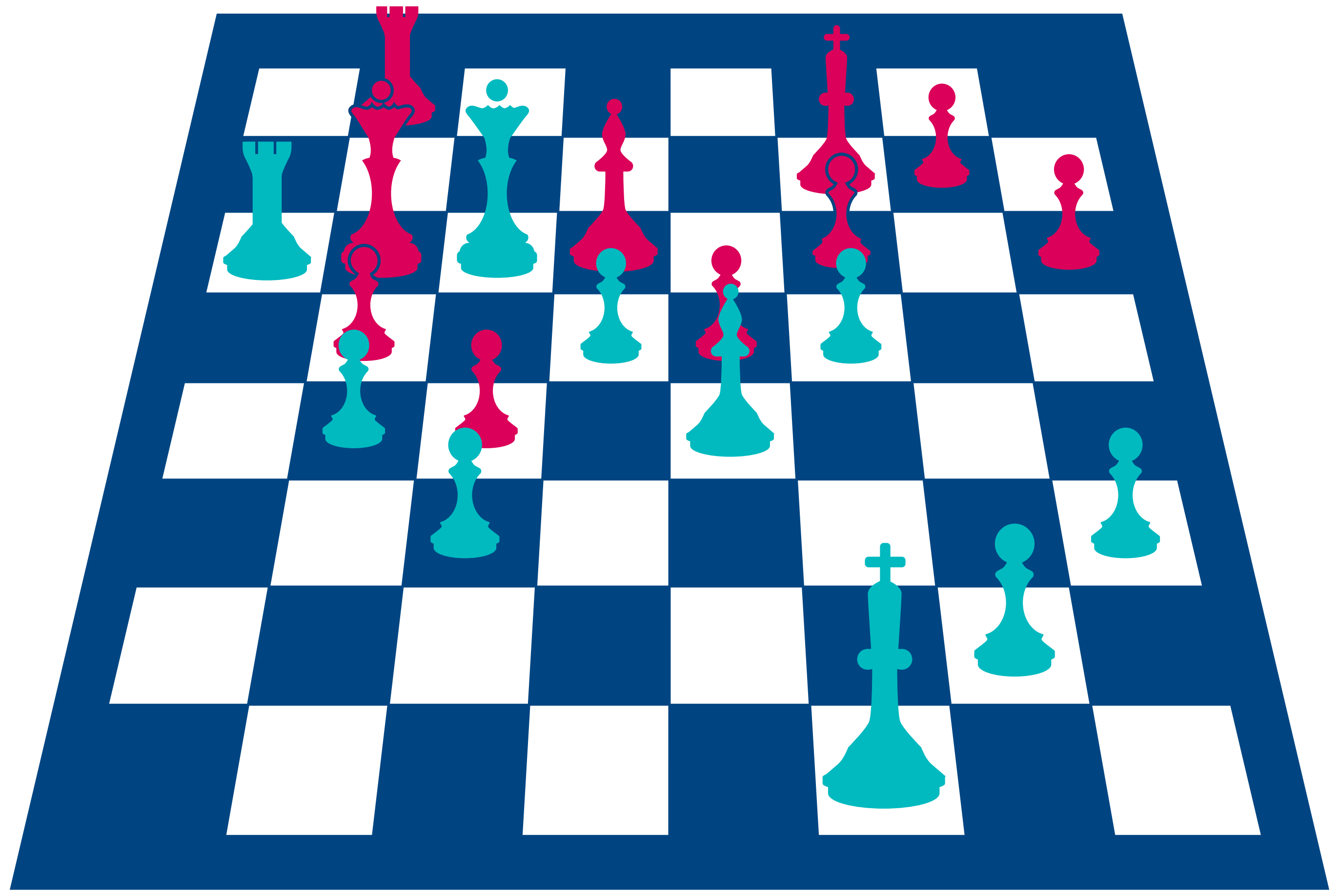
Ideal
Digital all-rounder

One wishful thinking for AI is that it can be used as a human-like friend and helper, letting us interact with the environment when we need it, for example as a voice-controlled assistant that starts the coffee machine in the morning, plans the route to work and prepares the shopping list for dinner.



Sci-Fi
Terminator

In literature and film, AI is often embodied as an independent and autonomous being, for example in the film "Terminator," where AI is cast as a threatening human-like machine with a mind of its own. This image deliberately stokes our fear that, in the battle of man against machine, the machine would win.



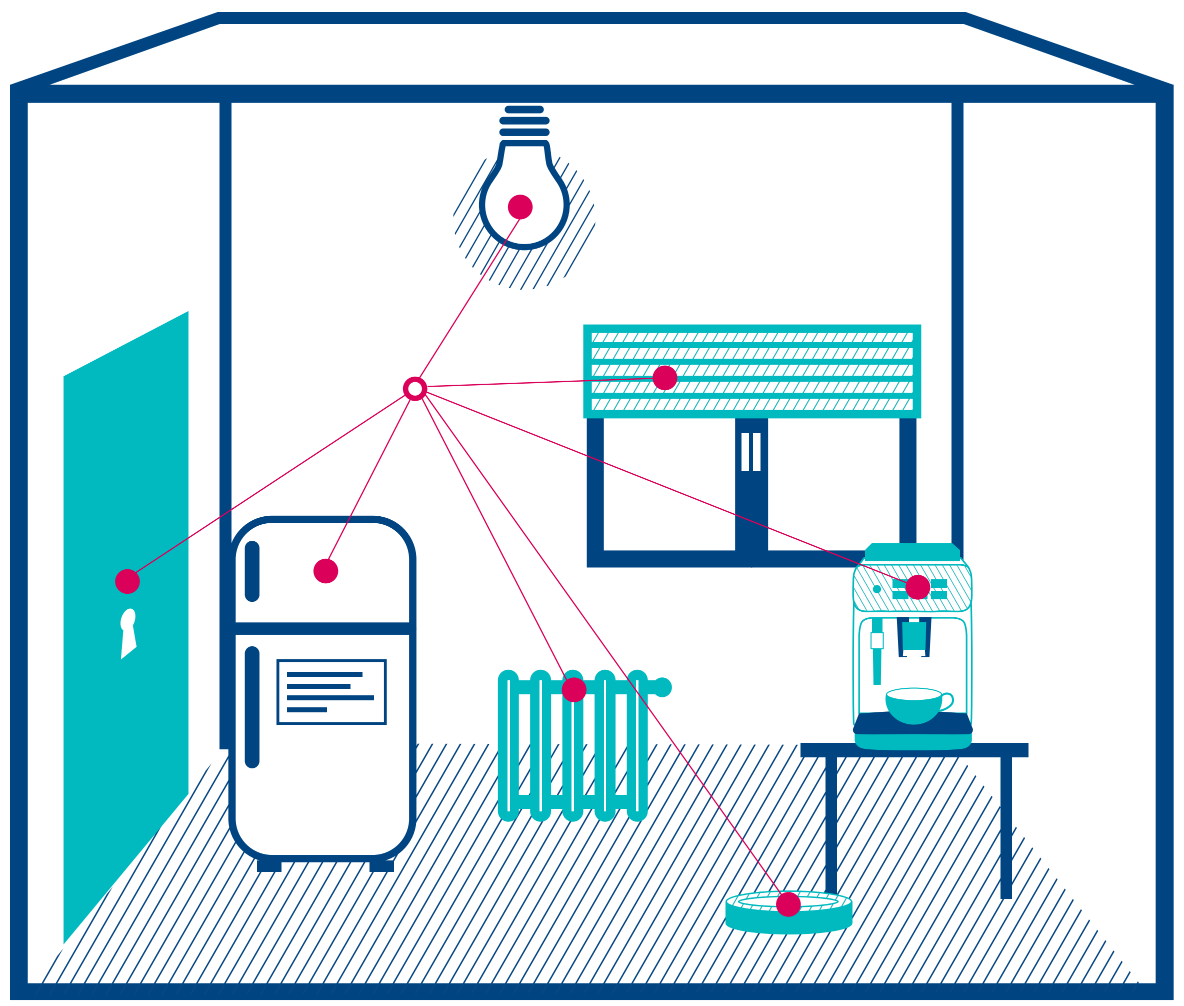
Reality
IBM's Deep Blue

In reality, AI is a computer program that can solve complicated problems quickly and efficiently by processing large amounts of data, such as IBM's Deep Blue. In the late 1990s, the chess AI defeated world chess champion Garry Kasparov – however, it needed human help.

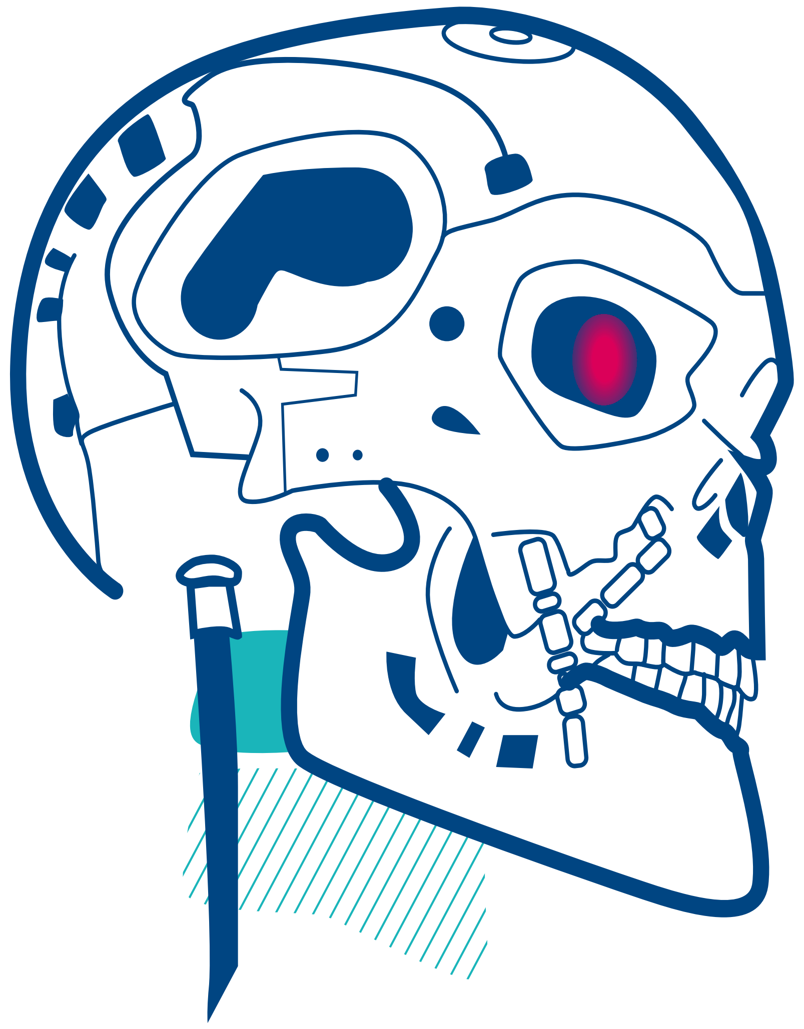
› Did you know?

› Menu

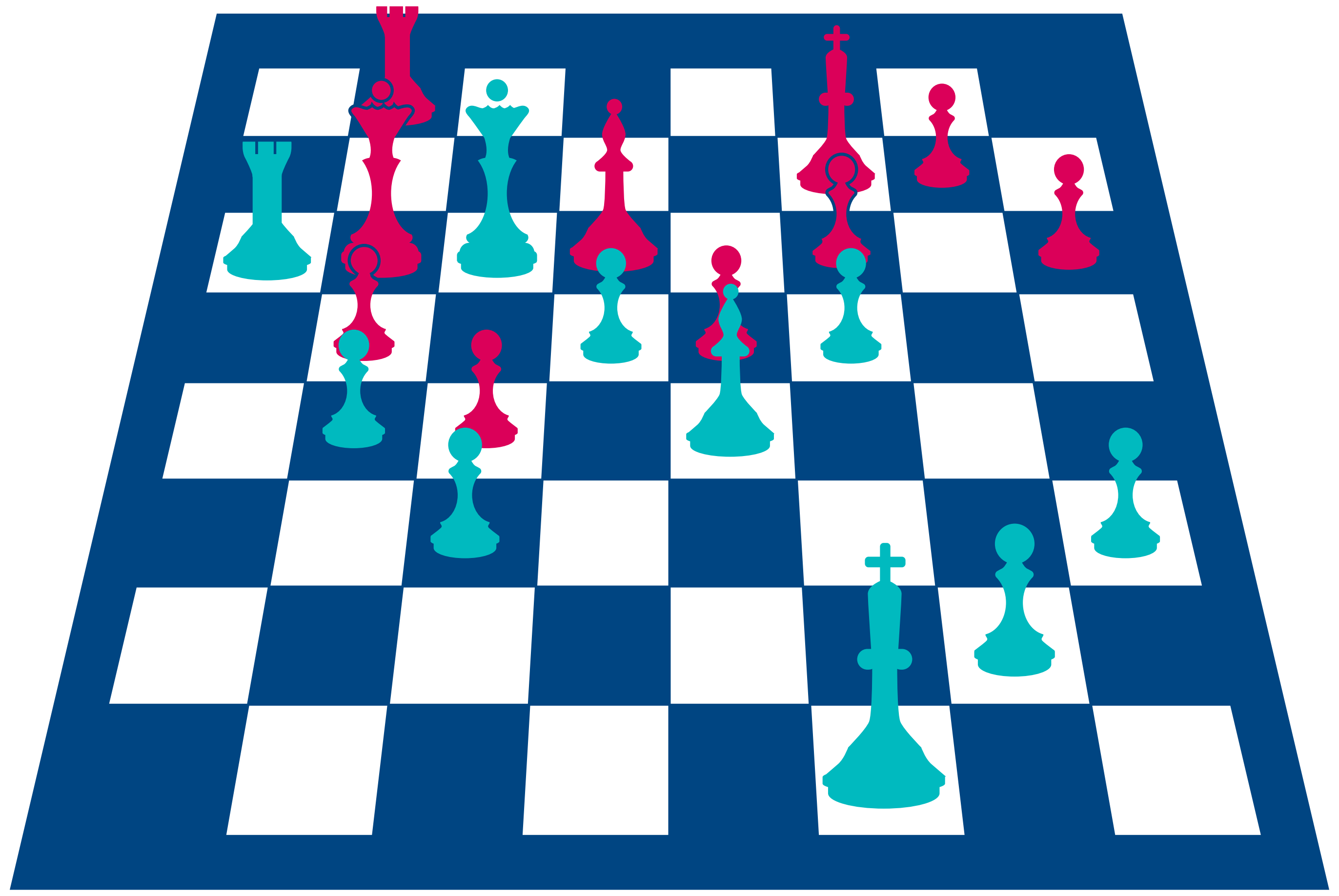
1.1 Fiction and reality



Ideal



Sci-Fi



Reality

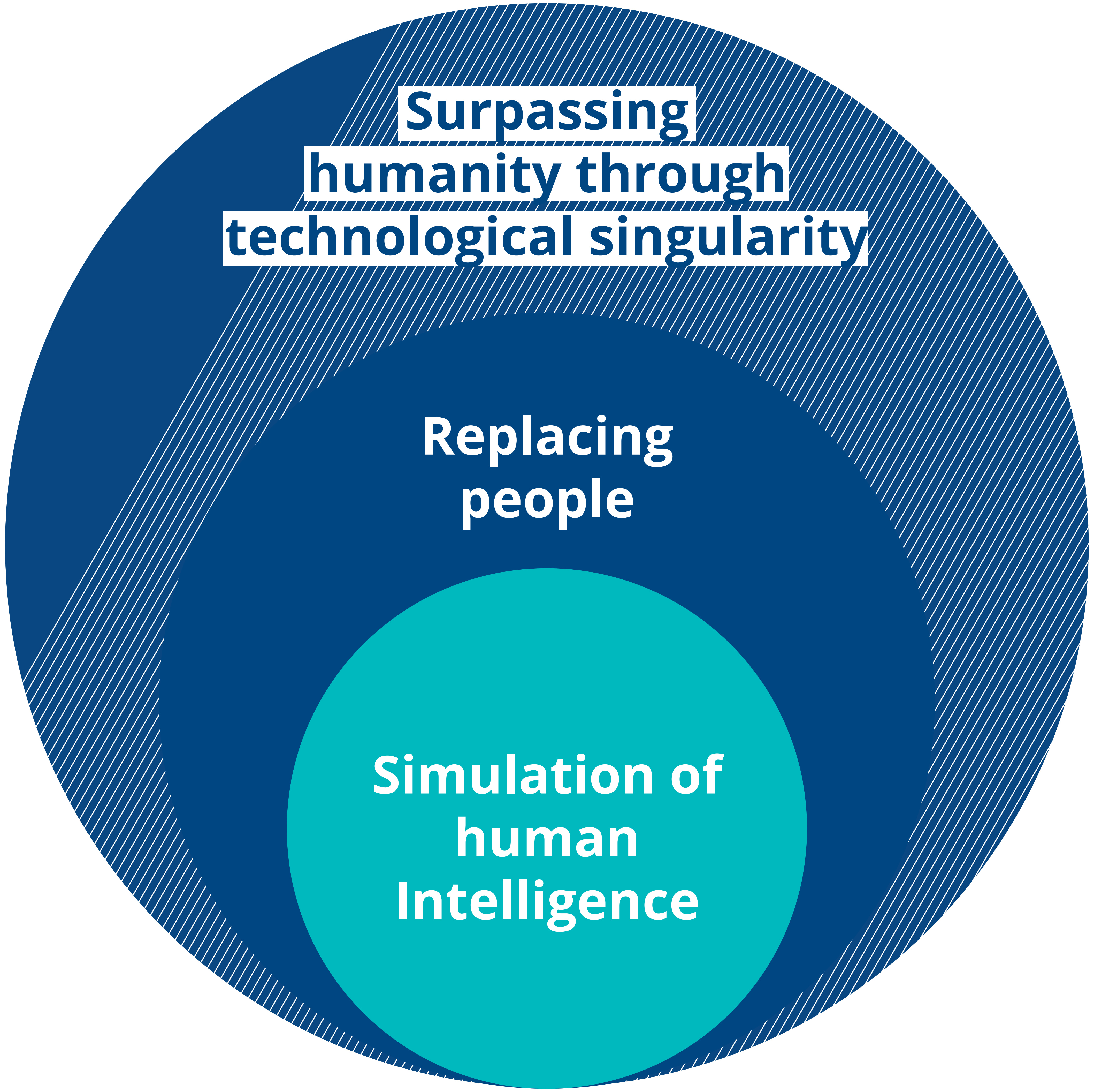
Strong AI and weak AI

Wishful and sci-fi portrayals of AI can be grouped as **strong AI**, which is characterised by the ability to solve many tasks of varying complexity.

In contrast, **weak AI** is programmed for a specific activity in our everyday life. The Terminator can ride a motorcycle, hold a bouquet of flowers and have a conversation – Deep Blue, in contrast, can only play chess.

But how realistic is the development of a strong AI and how justified is the fear, that AI poses an existential threat to us humans?

1.2 Terminator fallacy



Reflections on strong AI, including the fear that it could wipe out humanity, stem from a fallacy, which can be explained by the theory of technological singularity.

The concept of **technological singularity** states that due to the rapid advancement of technology, AI will eventually surpass humanity in its capabilities. From this point in time on, it will continue to develop independently, meaning without human intervention.

This concept is based on the idea that human labour can be completely **replaced** by machines and automation. Machines can work faster and longer, and computers can perform computing tasks more precisely and efficiently than humans.

This idea follows the belief that human intelligence can be **100 percent simulated**. However, there are arguments that counter these claims.

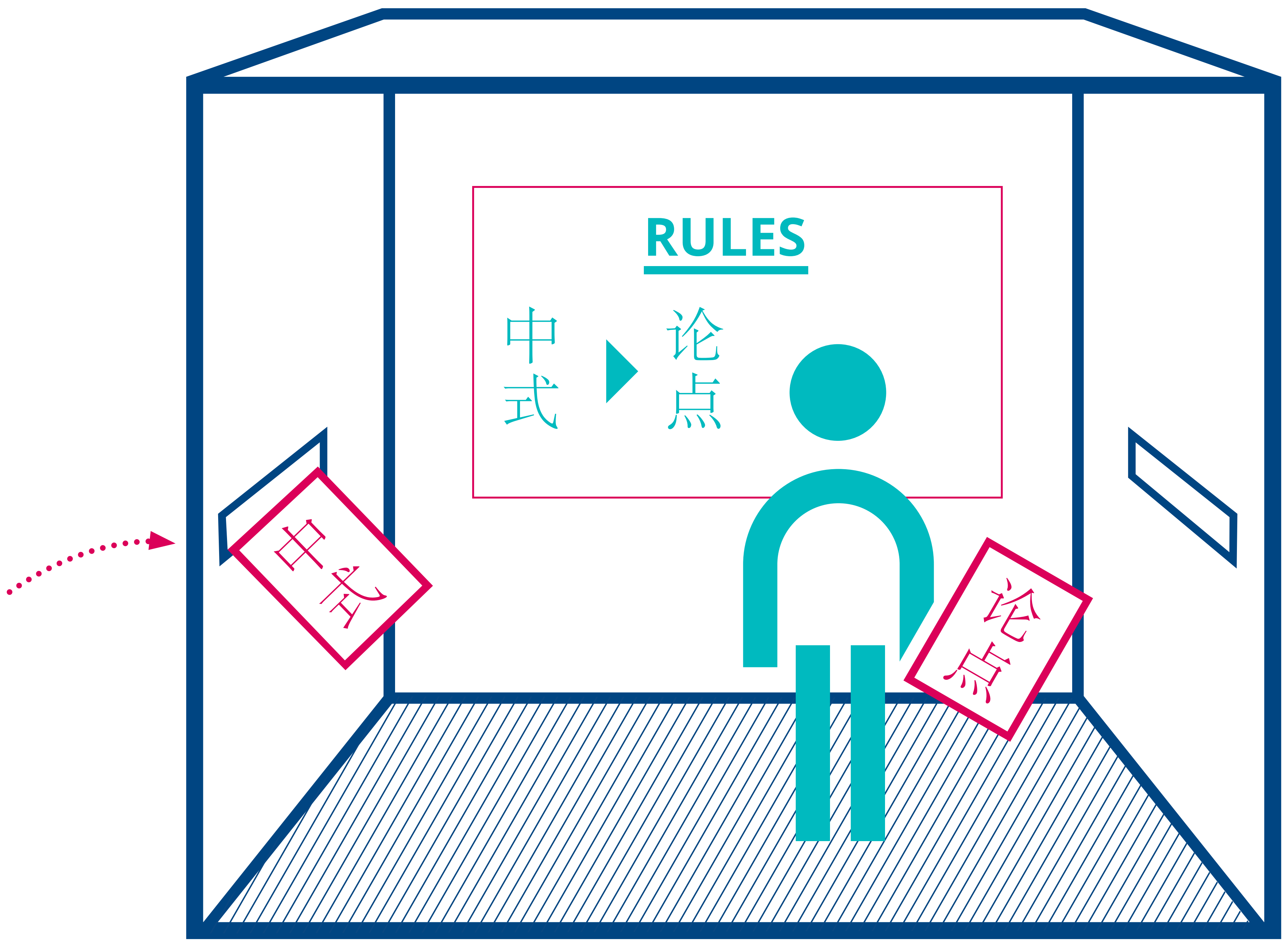
› Menu

1.3 Chinese Room Argument

The experiment

As early as the 1980s, the philosopher John Searle proposed a thought experiment that addressed the question of whether human intelligence can be simulated. He asked us to imagine the following:

1. We are sitting in a room, where we receive sheets of papers with Chinese characters through a slit in the wall. It is important to note that we neither understand nor speak the language.
2. There are instruction manuals in the room explaining how to string the Chinese characters together.
3. We assemble the characters by following the instructions and deliver them through another slit in the wall.

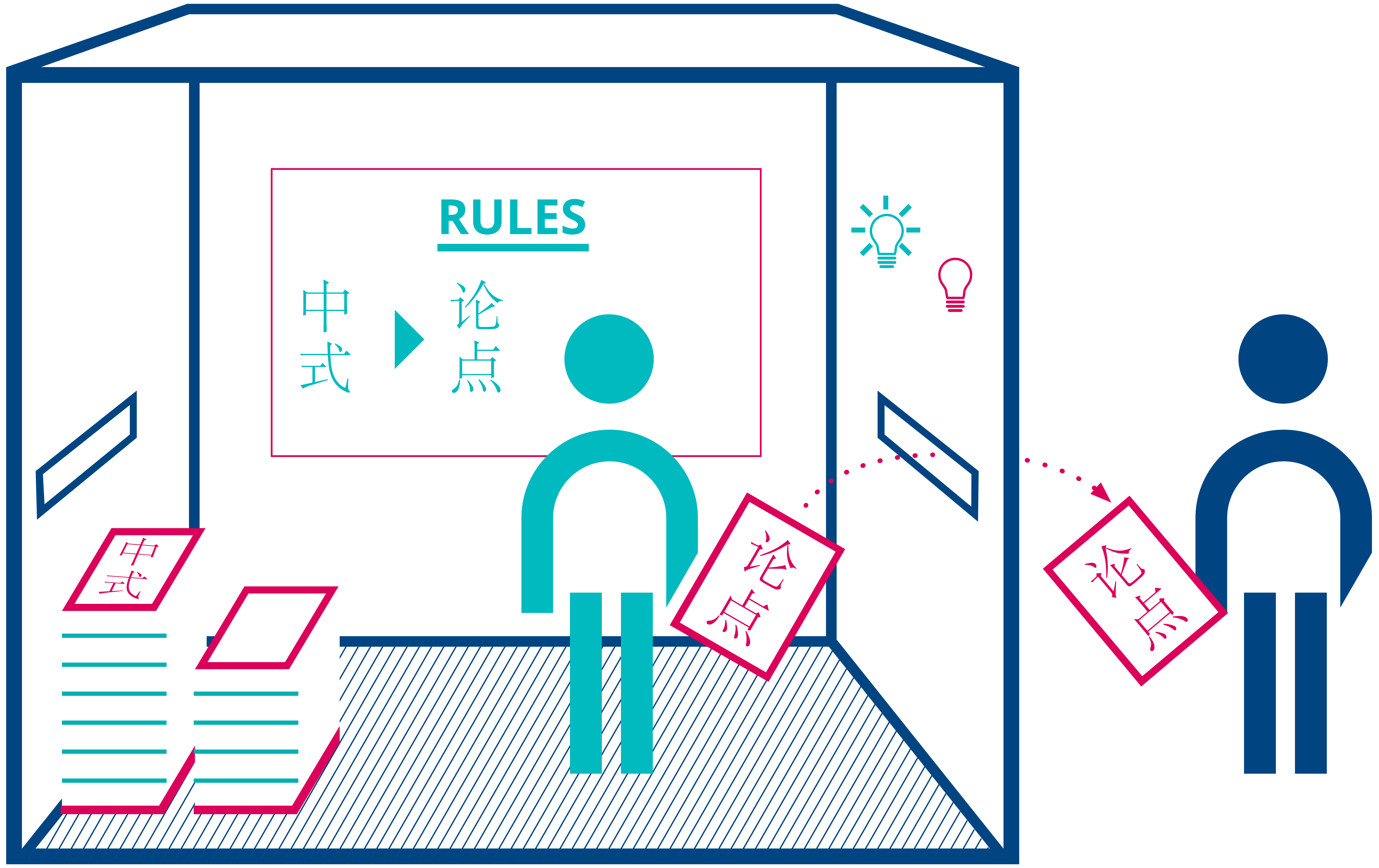


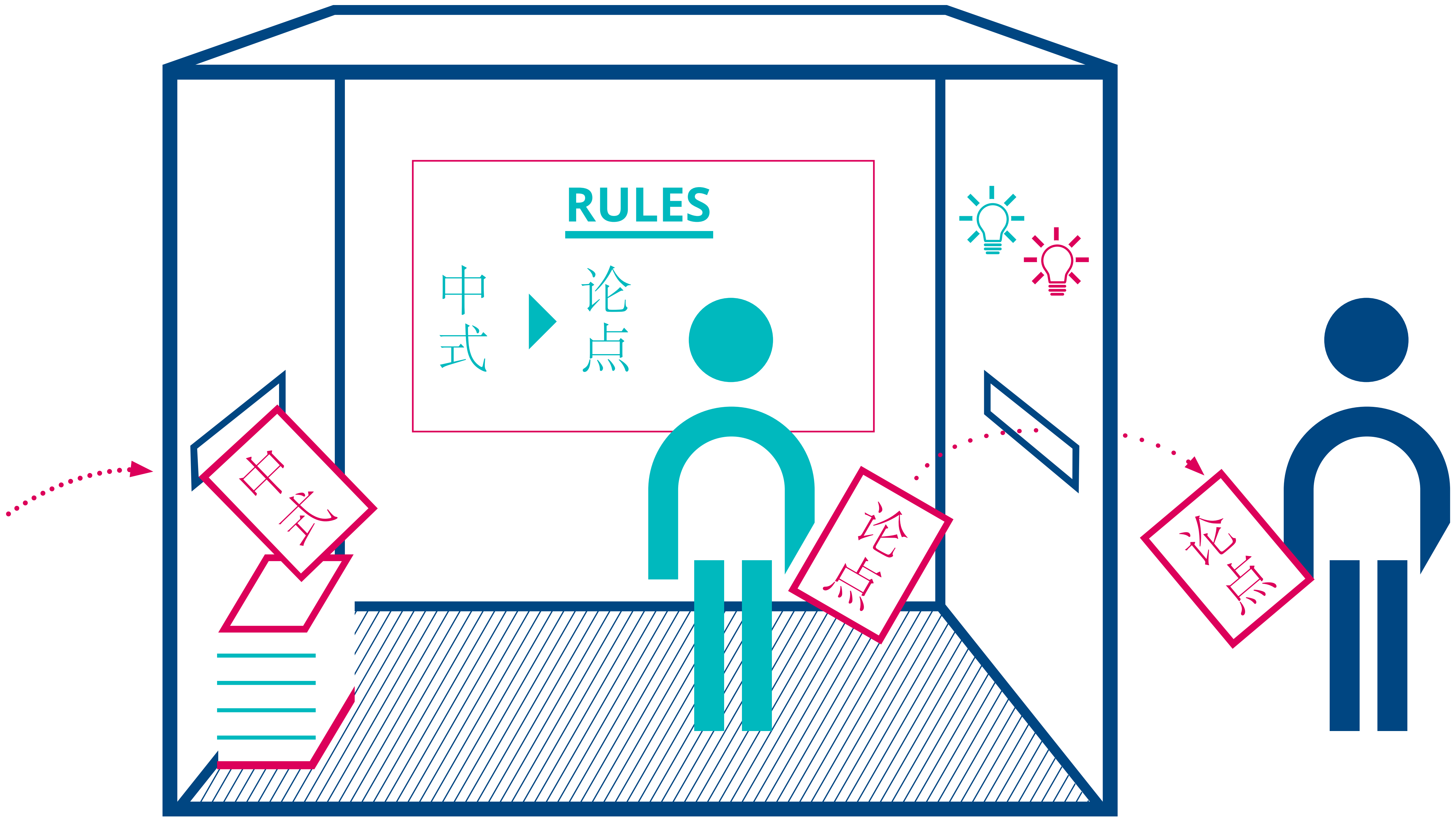
1.3 Chinese Room Argument

The effect

Outside of the room is a Chinese native speaker, who receives our sheets and evaluates the character sequences. The better and faster our compositions become, the higher the likelihood that the native speaker will have the following impressions:

1. Whoever produces the character sequences has mastered the language.
2. The smaller the lapse between input of characters and output of the response, the stronger the impression.
3. This impression will be reinforced by sequences of characters that are not contained in any rulebook, but are nevertheless correct. Like words of a dialect, they capture speech nuances that are reminiscent of native speakers.





Syntax does not suffice for semantics

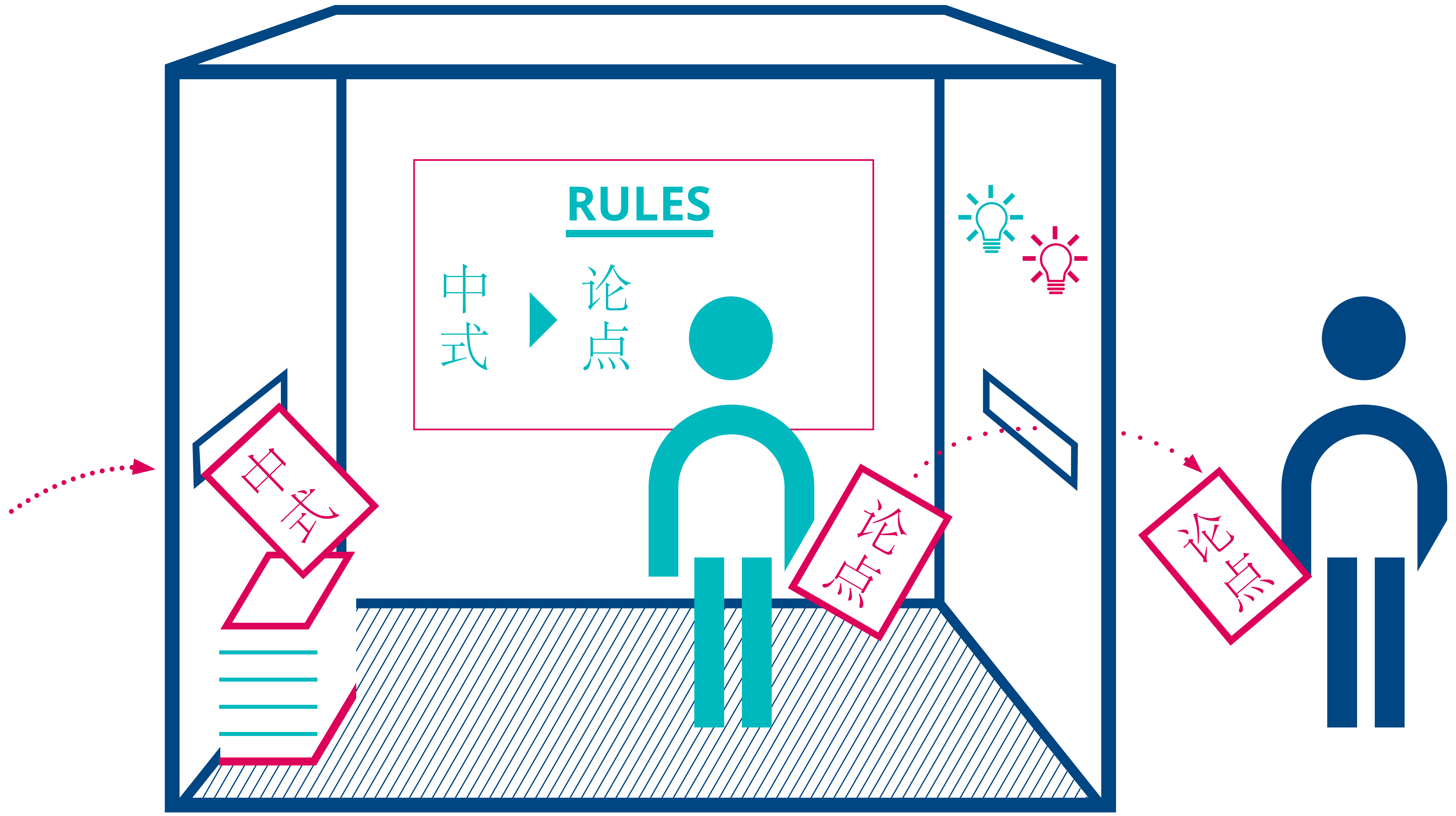
Conclusion

As long as our character sequences are correct, there is no reason for the native speaker outside the room to believe that we do not know the language. But the fact is:

1. We have no information about the meaning of the characters,
2. and therefore have no understanding about the message they are conveying to the person in their environment outside the room.
3. The characters have no significance for us, we only know the order in which they must be stringed together.
4. We don't speak the language, we can only use it.

John Searle concludes with the following insight: That something or someone is able to string symbols together correctly (syntax) is no evidence for the fact that this entity or person knows what these symbols mean (semantics).

1.3 Chinese Room Argument



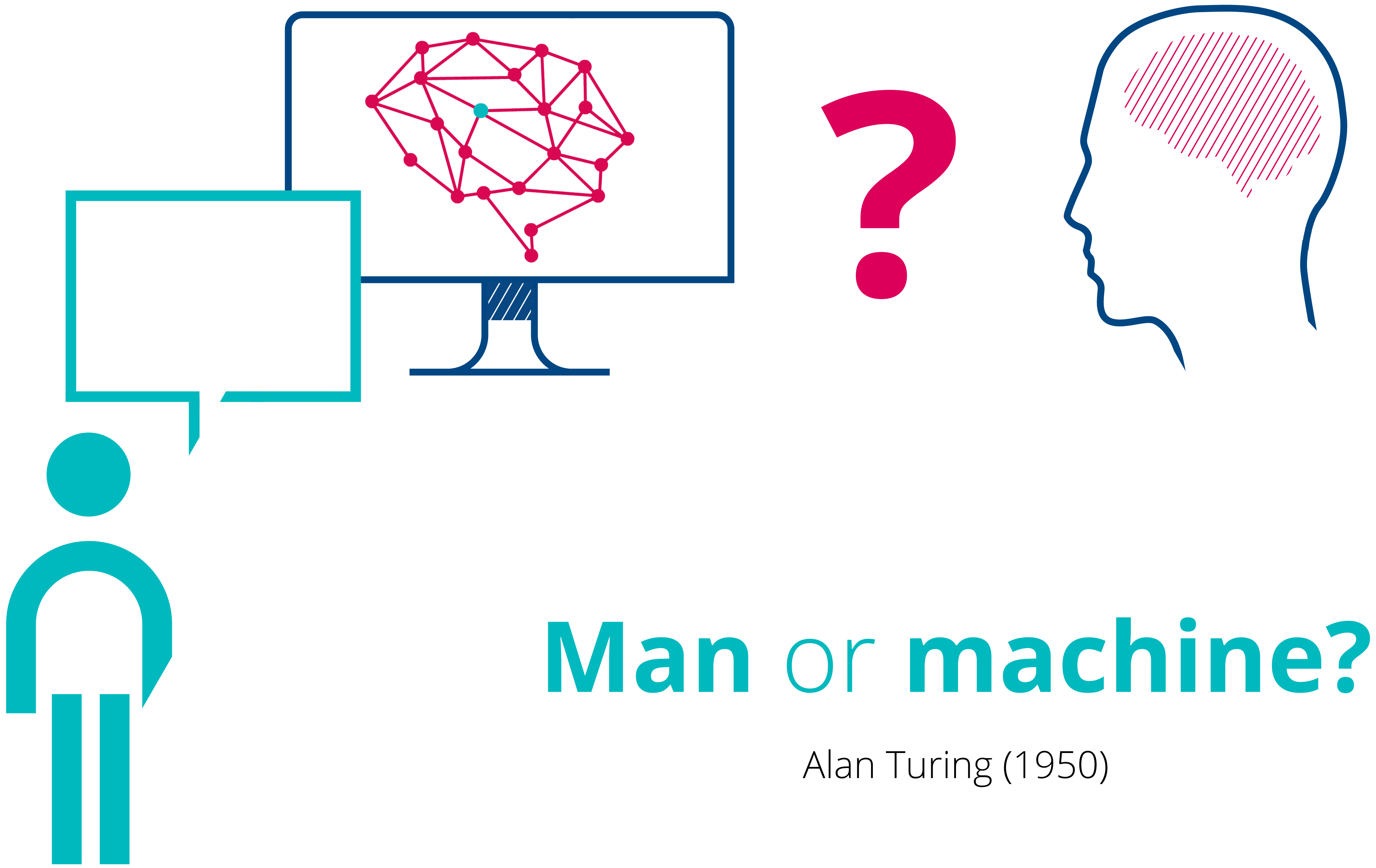
In summary:

Using John Searle's thought experiment, we can argue that computational power, or any other digital information processing, which is capable of simulating syntax, can at most be used to simulate human cognitive capacity, but it is not capable of thought itself.

This is a philosophical thought experiment whose conclusion we can agree or disagree with.

What about AI that are confusingly similar to humans?
Does it really make a difference whether a computer knows what it's saying?

Syntax does not suffice **for semantics**



Man or **machine?**

Alan Turing (1950)

Turing test

Alan Turing was a British mathematician who in 1950 came up with the idea of how to test an equivalent thinking ability between humans and computers.

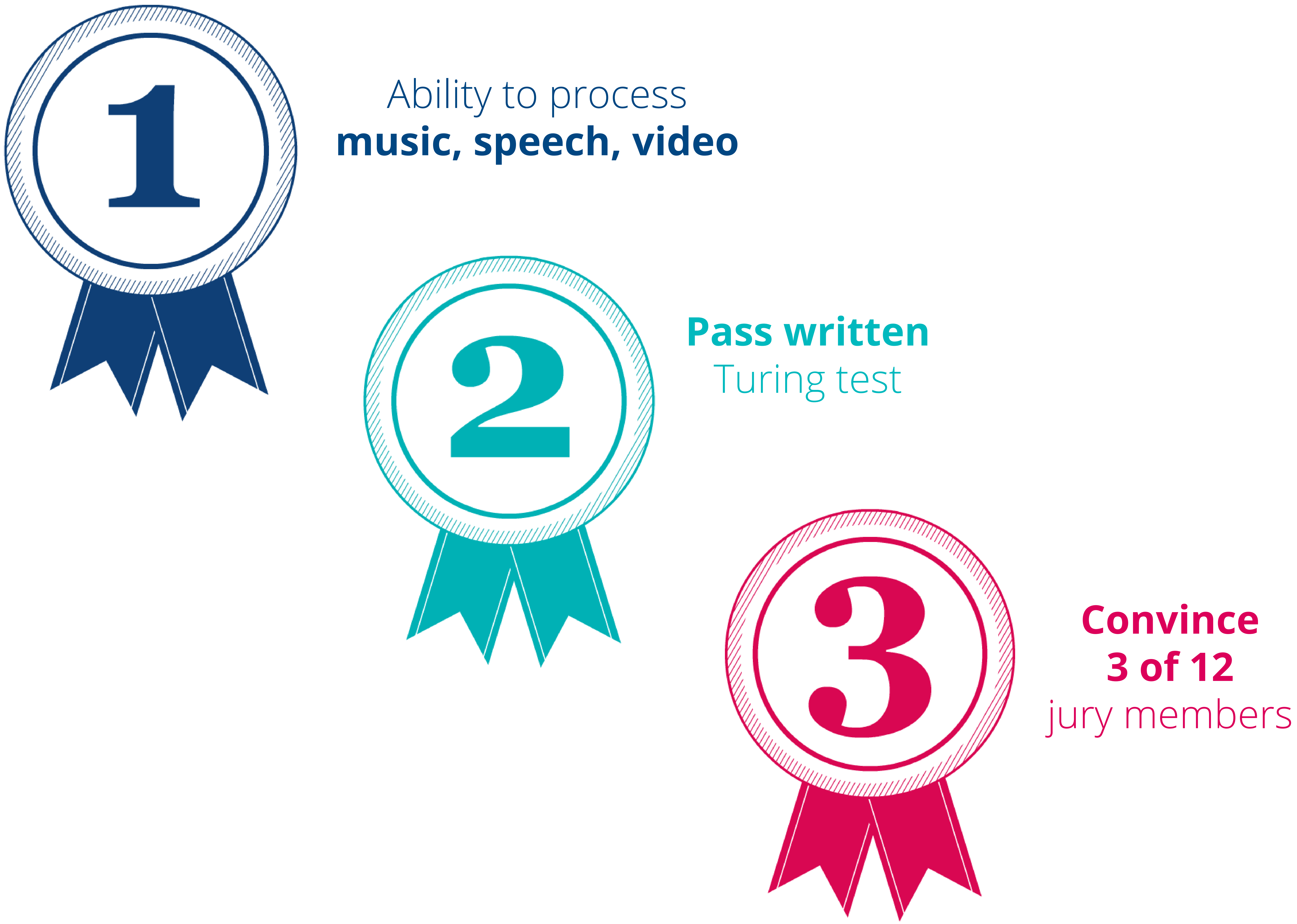
He developed a test (Turing test) in which a person communicates with two other actors. One of the two actors is a computer, the other a human being.

If the person cannot figure out through communication which of the actors is the computer and which one is the human, the computer has passed the test.

In the Turing test, it doesn't matter if the computer really understands what it is saying. If it communicates like a human, the computer can also think like a human, at least to the observer.

› Did you know?

1.4 Man and computer



Loebner Prize

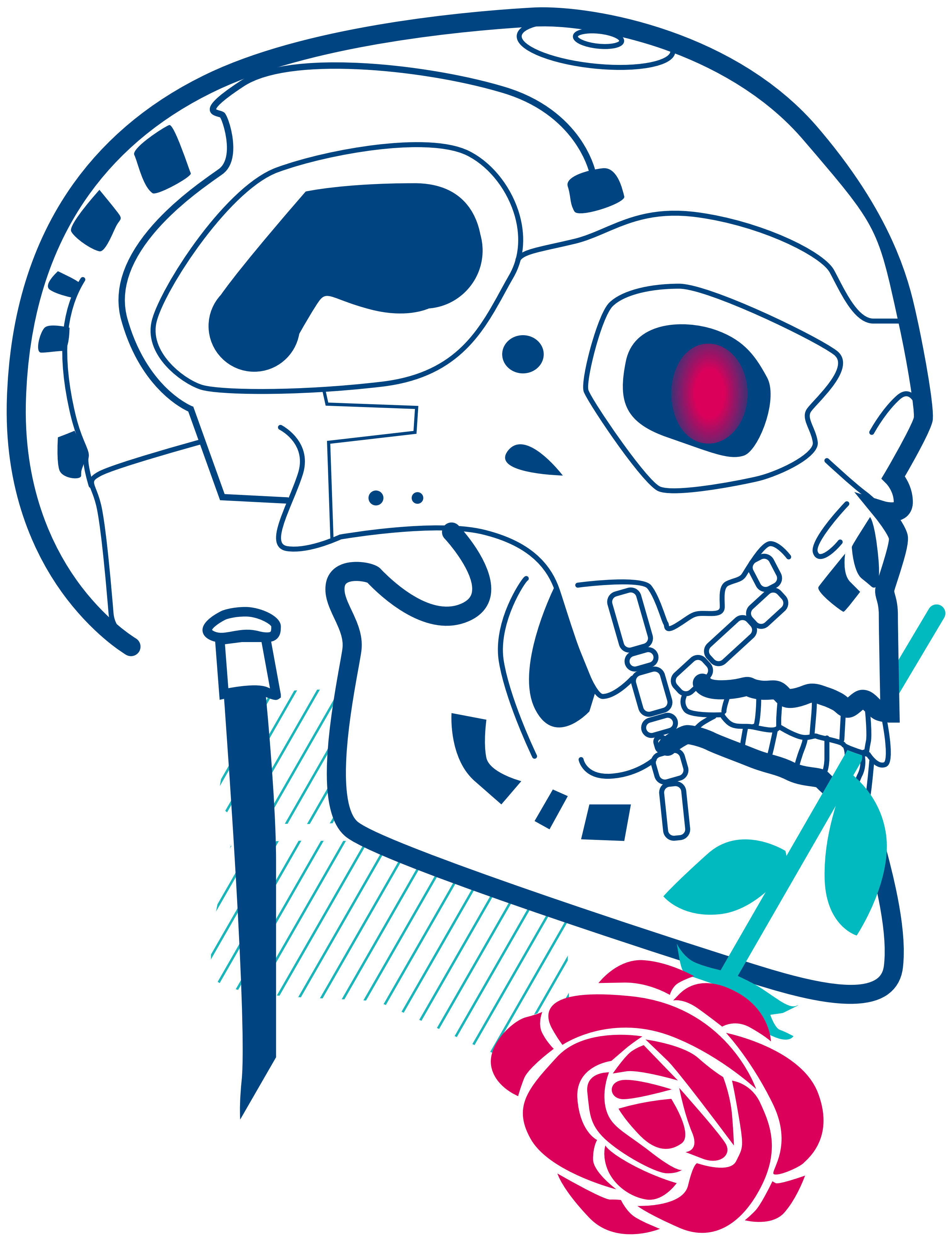
Since the early 1990s, the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB) has awarded the Loebner Prize. Participants in the competition submit the artificial intelligence they have developed with the goal of passing the Turing Test.

For bronze, the AI must achieve a certain score and convince a minimum number of jury members.

For silver, the AI must pass the written Turing test and convince at least half of the jury members.

For gold, the AI must pass the Turing test including the processing of music, speech and video to convince all jury members.

To this day, no AI has ever won more than the bronze medal.



Terminator fallacy

Back to the questions of why we don't need to be afraid of the Terminator and whether a computer can fully simulate human intelligence:

1. The Chinese Room Argument shows, using philosophical logic, that computers will never reflect or understand what they are doing or the meaning of their actions.
2. As a real-world example, the Loebner Prize shows that AI is still far from at least pretending to understand what it is doing.

In summary, there is no reason to believe that human intelligence will eventually be replaced by computers. The Terminator continues to only be part of our pop culture.

But what makes artificial intelligence intelligent and what is artificial intelligence anyway?

KONRAD ADENAUER STIFTUNG

2

What does "AI" stand for?

What does "AI" stand for?

To understand what makes artificial intelligence intelligent, it helps to take a closer look at the meaning of the word itself. Let's see what the word means:

2.1 The “A” in AI

Artificial

The word artificial refers to the nature of AI, and two aspects that fall under it:

AI is never organic and always embedded in some type of computer program. AI is therefore a software that we can‘t see and experience until it‘s integrated into a piece of hardware.

AI is created for a specific purpose. It is not created out of a whim of nature, but programmed to solve a particular problem.

In the case of autonomous driving cars, the car itself is not AI, only its software, which was programmed to solve the problem of navigating a car from A to B without human intervention.

2.2 The "I" in AI

Intelligence

One of the biggest problems in understanding artificial intelligence lies in the word intelligence. The association with human intelligence is obvious.

In the case of AI, intelligence means solving a new problem based on collected information without any outside help. The process of finding a solution consists of three aspects:

› Data processing

› Efficiency improvement

› Autonomous & automated action

2.2 The “I” in AI

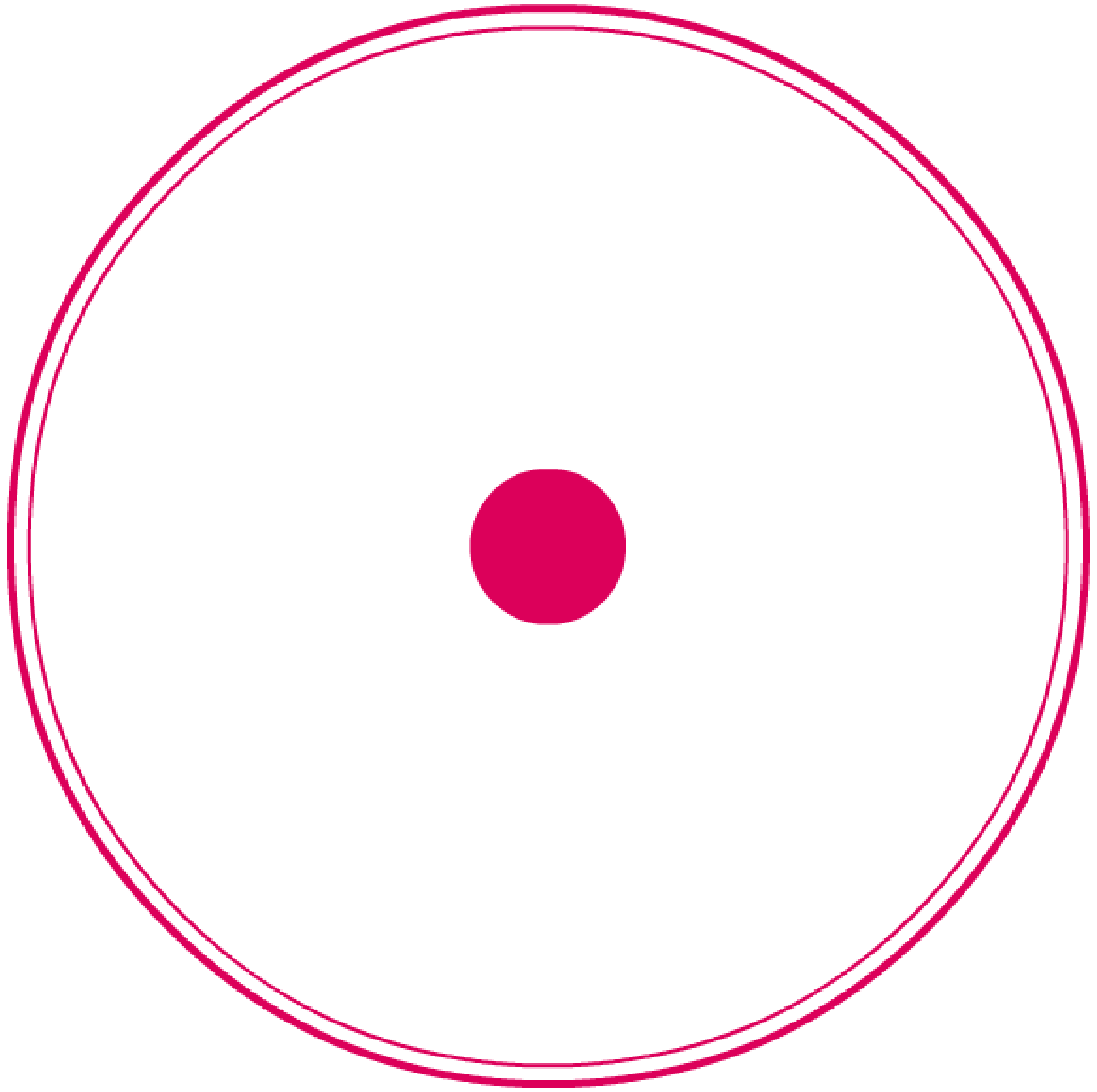
INTELLIGENCE

Data processing

In the process of solving problems, AI processes data representing the human environment. This environment is not perceived by AI the way we humans perceive ours, but

1. AI translates the environment into processable data via sensors that have been developed and selected by humans
2. and/or the AI receives filtered data directly fed in by humans.

When it comes to data processing, humans remain the primary influence on the AI’s approach to problem solving and ultimately also responsible for the outcome.



Data processing ≠ perception of the environment

› Back
› Menu

2.2 The “I” in AI

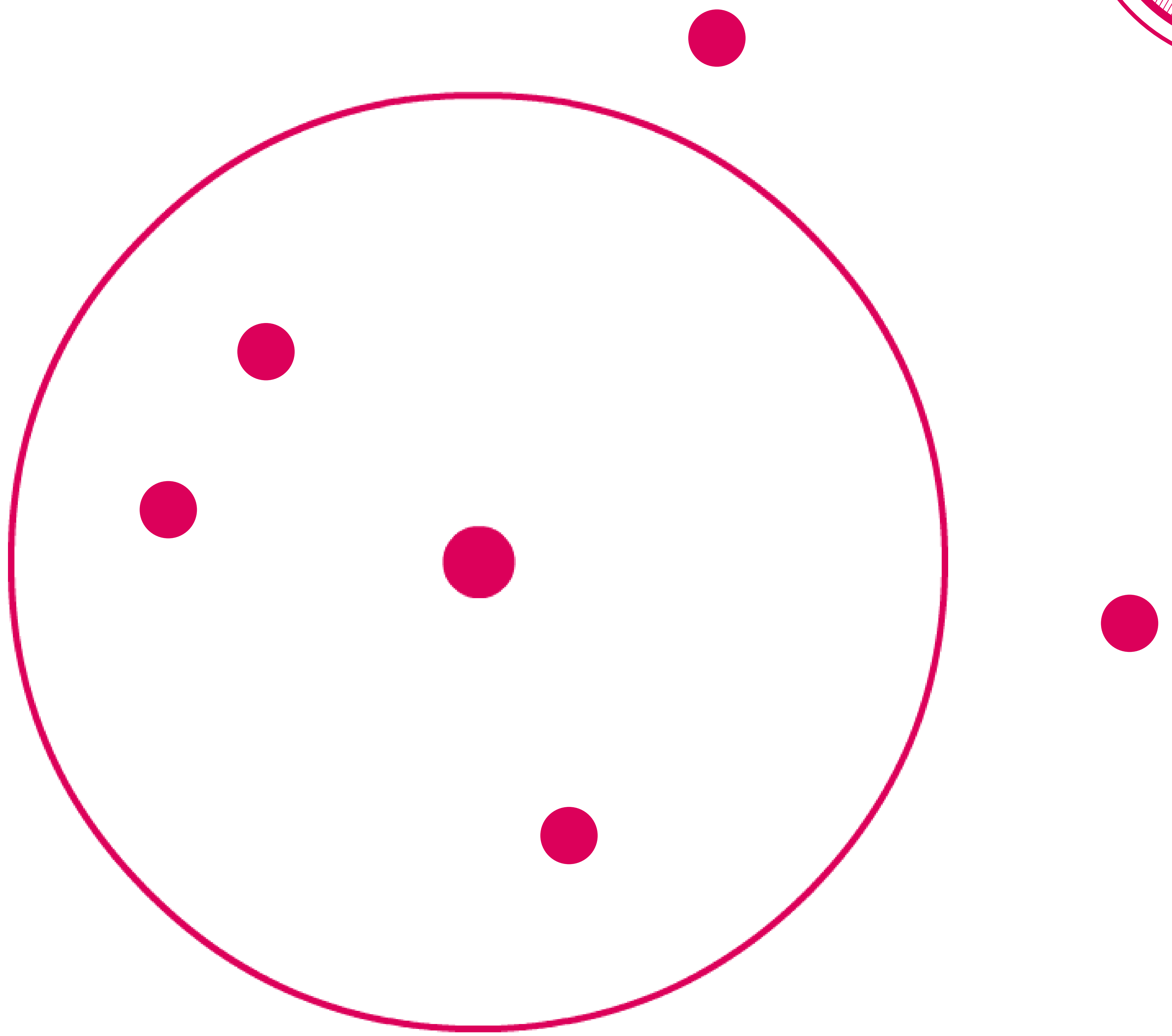
Efficiency increase

AI can produce its own data and use it to optimize new approaches to solutions, if this function is intended by humans. The underlying process is

1. an algorithm driven by computing power. In other words, a pre-programmed protocol that defines which rules must be followed when processing the data,
2. recursive, i.e. the rules are applied to the solution created by the rules themselves.

Within this process loop, humans must determine which new proposed solutions are “right” or “wrong.” AI does not learn new solutions based on experience, but processes data using a human-manipulated trial-and-error process to build up successive solutions.

› Did you know?



Efficiency increase ≠ learning

2.2 The "I" in AI

INTELLIGENCE

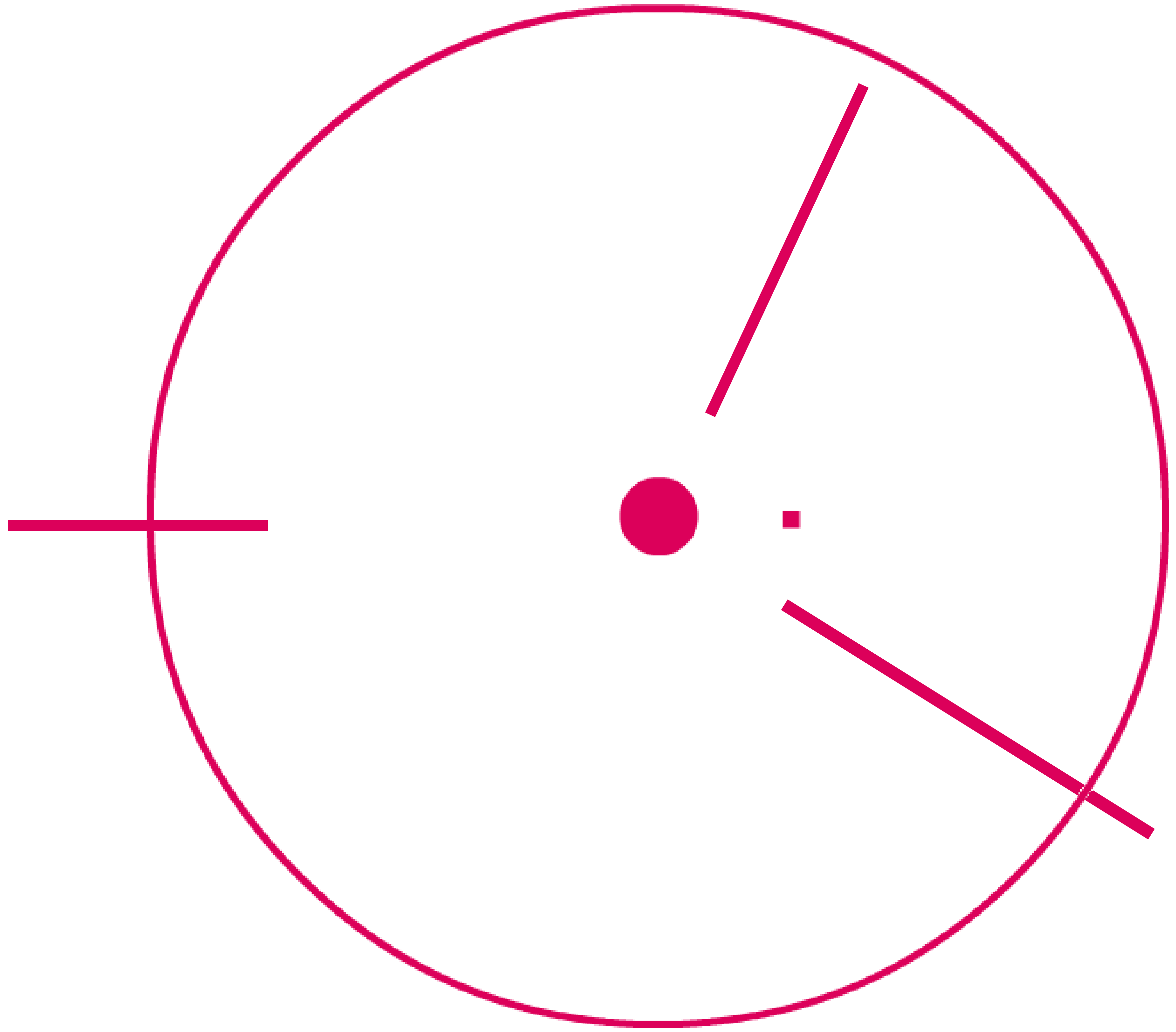
Autonomous & automated

AI can work (partially) autonomously as well as automated. In both cases, the AI reaches the goal without human assistance. In autonomous processes, however, the impression can arise that the AI makes an independent decision, according to its own sensibilities.

Since the number of causal relationships that can be programmed into an algorithm is limited, AI can sometimes come up with suprising solutions that were not intended by the human programmer. However, this has nothing to do with self-determination.

AI does not make a decision according to its own sensibilities, but offers solutions that are as close as possible to human specifications.

› Did you know?



Autonomous & automated ≠ self-determination

› Back
› Menu

KONRAD ADENAUER STIFTUNG

3

Summary

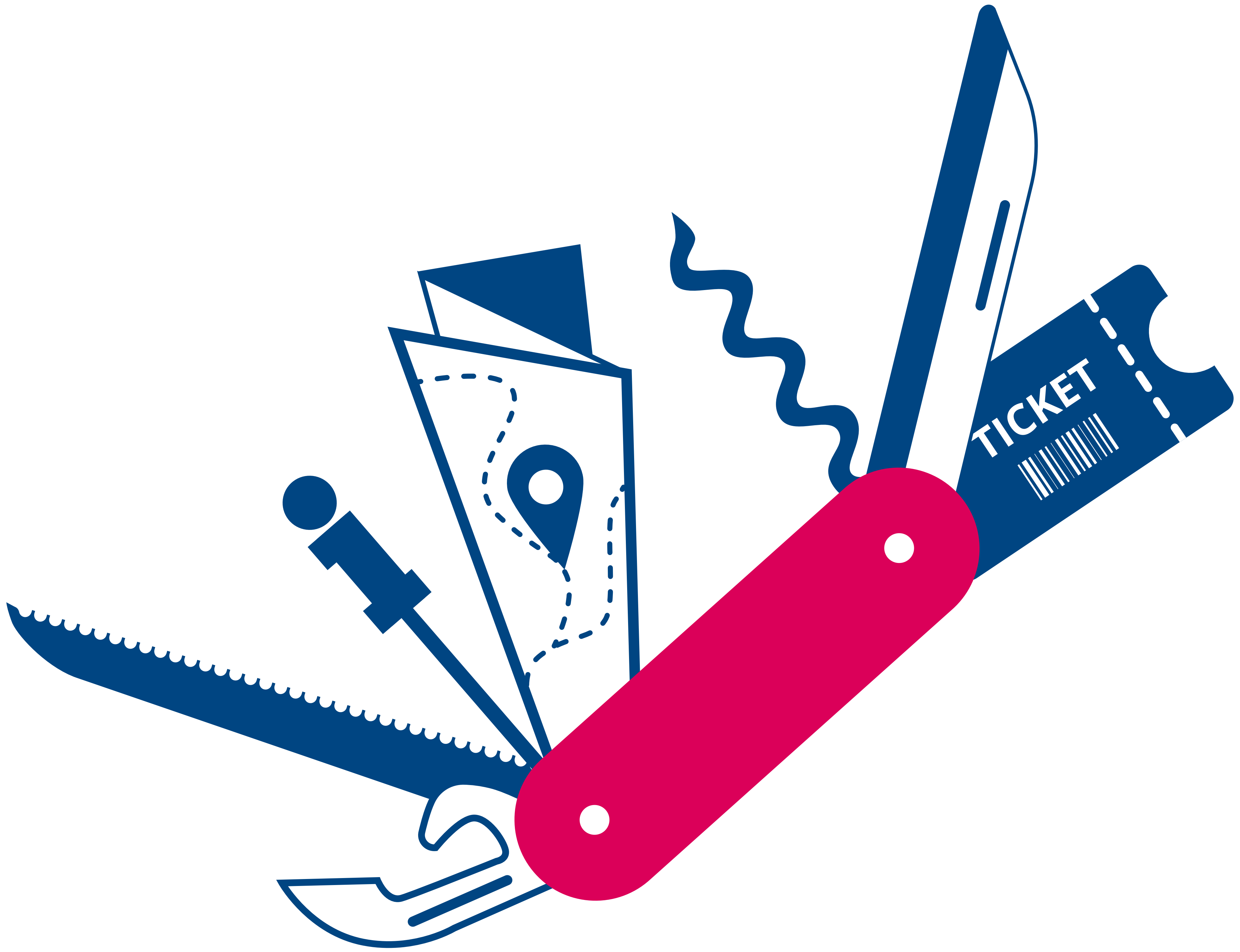
3.1 AI as a tool

Finally, let's take a step back and look at artificial intelligence as a tool.

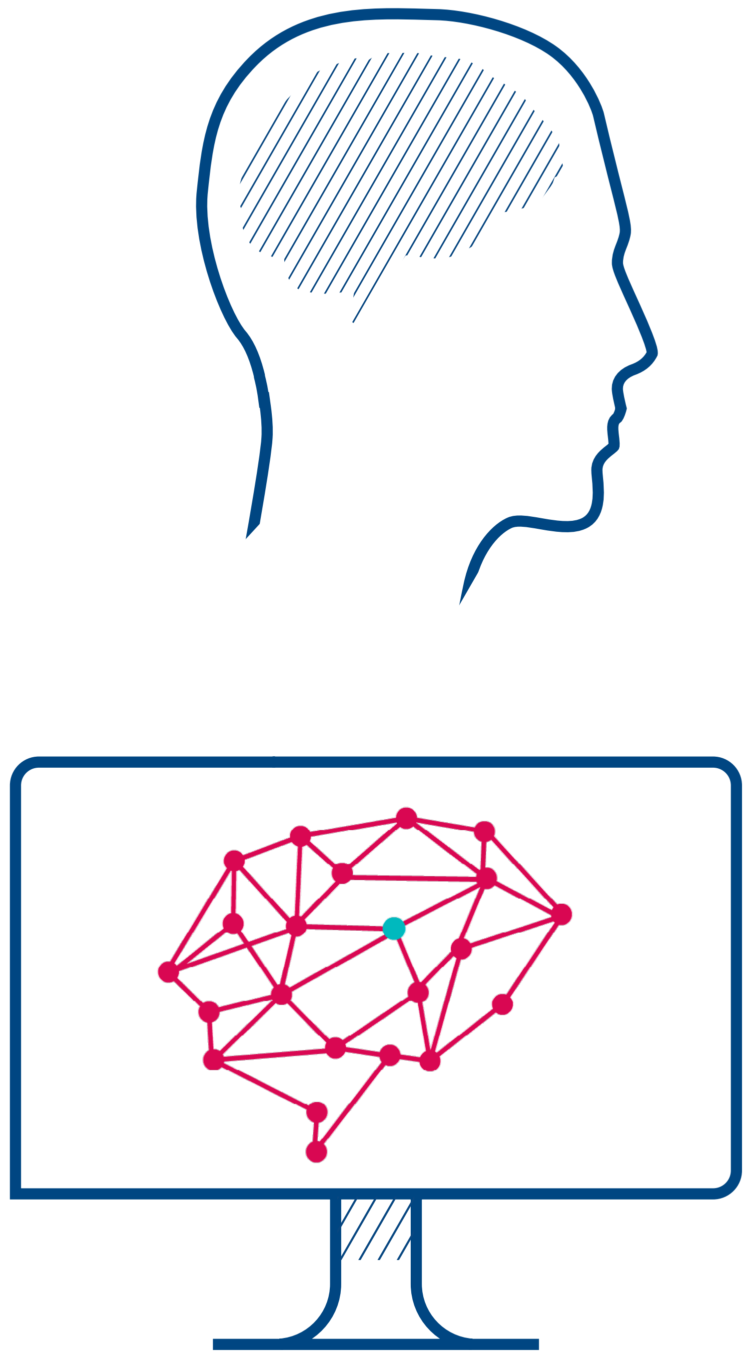
Let's assume that you are in a foreign city and want to visit a popular restaurant by public transport as quickly and conveniently as possible.

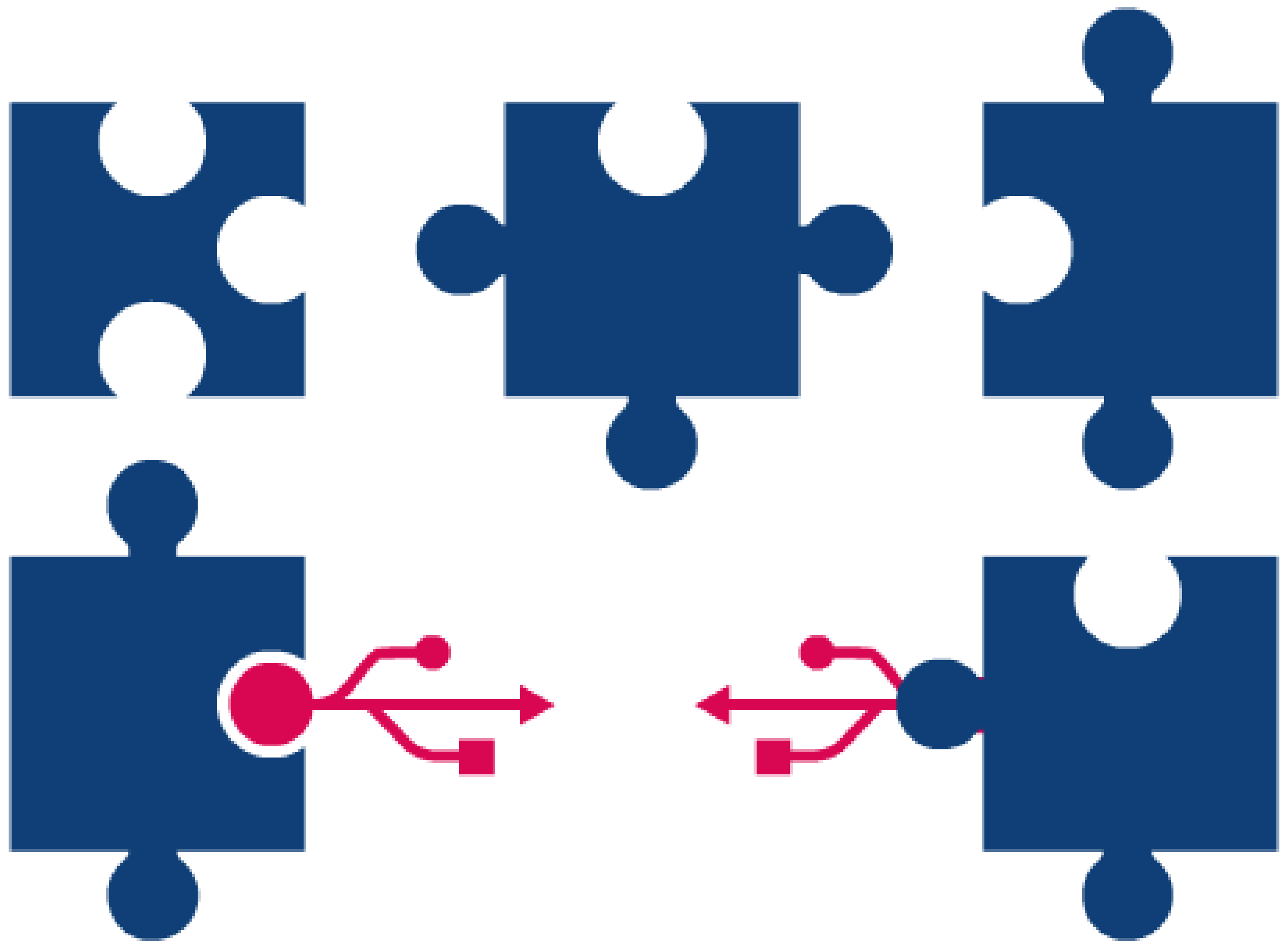
Whereas at the beginning of the millennium you would have needed a good travel guide, a ticket for public transport services and a nice chat at the tourist information point, 20 years later the search process is done via your smartphone.

Like a sort of Swiss Army knife of the digital space, AI serves as the underlying general-purpose tool. It offers suitable solutions for different problems. However, the decisive factor is not only AI. Equally important is the Internet as the world's biggest library.



3.1 AI as a tool





Due to the increasing conflation of the analogue and digital worlds, we as individuals are dependent on reacting adequately to this data and information overload.

AI can help to separate the unimportant from the important information and offer an individually tailored solution from the mass of different data available.

Similar to the way we systematically analyse all the individual pieces for fit when looking for the right puzzle piece for a problem, we can use AI to find the right puzzle piece for us. Seen in this context, AI is a method for recognising data patterns that help us solve a problem.

AI is a method of
recognising patterns from data

3.2 Epilogue

In the previous slides we have seen what Artificial Intelligence is and what it is not.

There is no reason for fear – neither on a philosophical level nor in practice:
AI cannot replace humans.

AI is not a digital brain that thinks better and faster than we do. AI is a digital tool that helps us transfer our human logic (or human thinking) – which is ultimately our categorisation of the analogue world – to its digital copy.

At the center of AI is the human being, with all their strengths and weaknesses.

In the next series, we will look at the opportunities and risks involved in dealing with artificial intelligence.

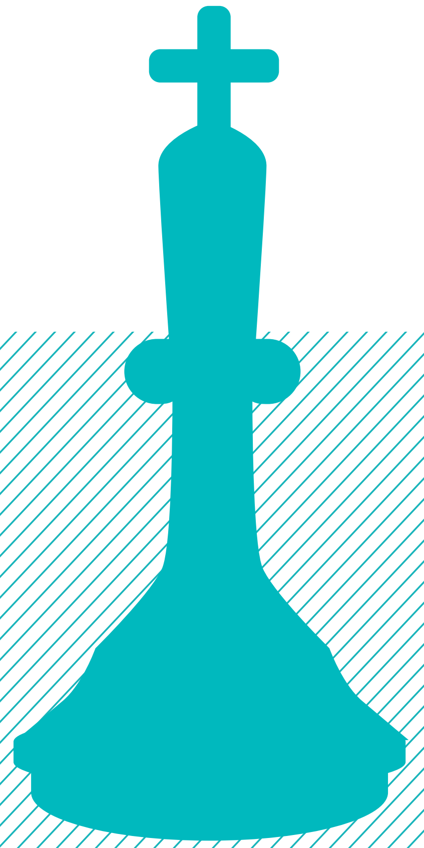
In doing so, we would also like to explore the question of why the use of AI is changing the social role of humans in society and what the use of AI has to do with fostering human creativity.

› Menu

Appendix: Did you already know?

Did you know?

AI is not perfect



1.000.000.000.000.000.
000.000.000.000.000.
000.000.000.000.000.
000.000.000.000.000.
000.000.000.000.000.
000.000.000.000.000.
000.000.000.000.000.
000.000.000.000.000.

Kasparov

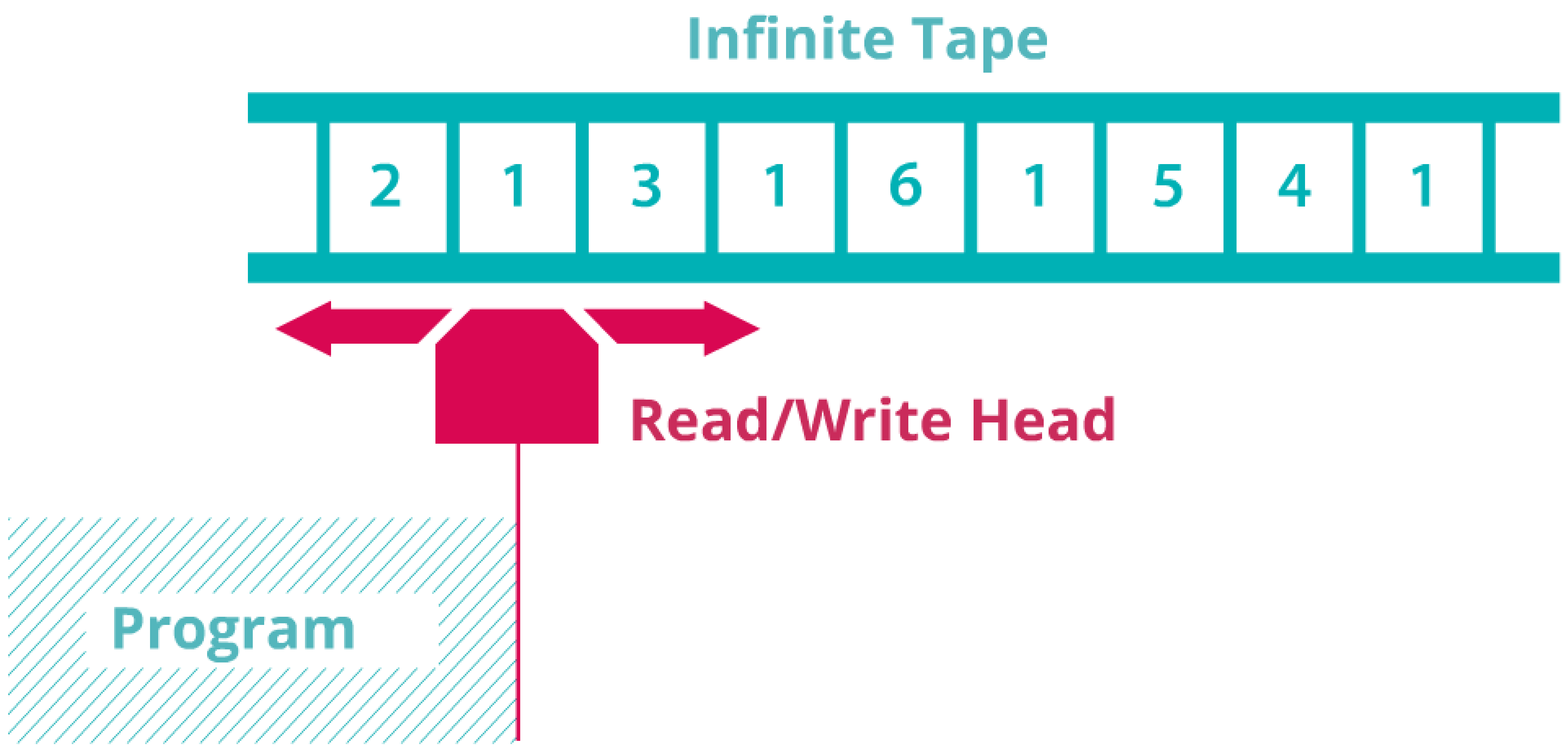
That an AI can be superior to humans was demonstrated when IBM's chess AI **Deep Blue** defeated chess grandmaster Garri Kasparov in a competition. However, the "AI defeats humans" narrative is distorted because IBM's engineers actively intervened. Not only did Deep Blue have all the information on Kasparov's playing behaviour; between chess games, IBM was allowed to customise Deep Blue's software and fix bugs in the source code.

Shannon number

But how does an AI actually play chess? AI uses the so-called brute force approach to play chess. It calculates every possible course of the game, which moves are possible, which reactions of the opponent are possible and so on. In theory, chess AI can generate the complete game tree. You therefore cannot set a trap for it; your own mistakes, in contrast, are punished immediately.

However, the number of possible moves is so high that a concrete figure cannot be calculated. Mathematician Claude Shannon has provided an estimate, known as the Shannon number, which states that a chess game has **10^{120} possible** game variations. A modern high-performance computer needs **10^{90} years** for this, i.e. longer than the existence of our universe: the time since the Big Bang multiplied by one billion. Therefore, a chess AI will most likely never be able to play chess perfectly.

Did you know?
Alan Turing's Turing Machine



Alan Turing

Alan Mathison Turing was a British mathematician and is considered to be the founder of modern computer science and an influential pioneer of artificial intelligence. The reason for this is the so-called **Turing machine.** A machine existing only on paper, to be understood as a theoretical object in mathematics. Although it does not physically exist, the Turing machine still represents the benchmark of modern computer performance.

No computer in the world can do more than the Turing machine, yet its operation is simple. With his machine, Alan Turing pursued the goal of making the term **algorithm** tangible. In as simple a way as possible, he wanted to represent how a machine follows rules to produce a certain result. It is based on a tape divided into cells with numbers. Next to the tape is a read and write head that can recognize the numbers and rewrite them to another number if needed.

In this head there is an algorithm that shows exactly what must be done as an action for which recognized number. Example: If the head recognizes a 6, it should erase the 6, write a 2 and move three cells to the right. Depending on the pre-programming of the algorithm, the process can progress at will until a desired result is achieved. All modern computers work according to this principle, whereby the underlying programs do not work with arbitrary numbers, but with 0 and 1.

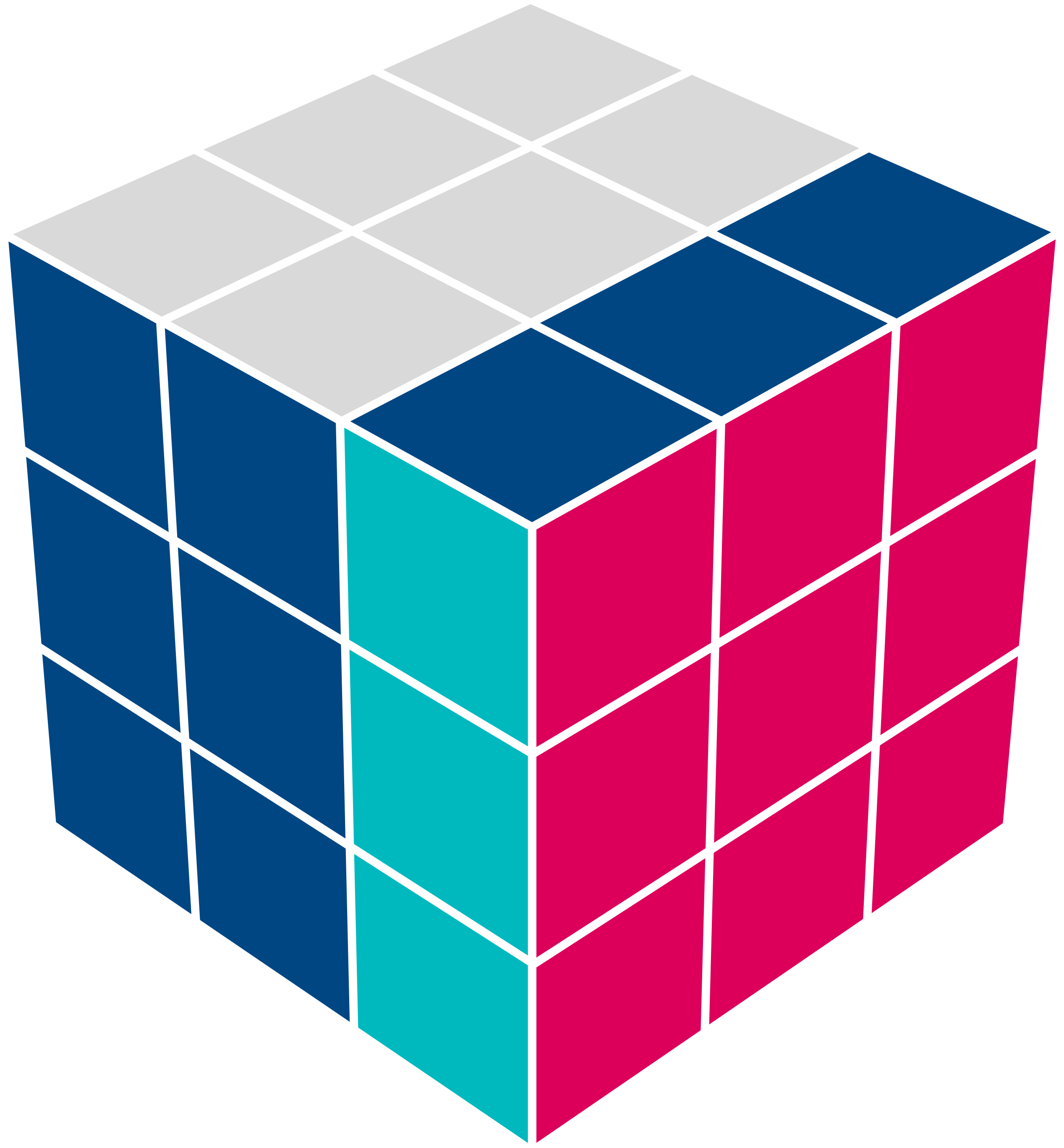
› Back
› Menu

Did you know?
Algorithm

Simply described, an algorithm is a set of instructions that defines the sequence of a task, and therefore provides a step-by-step procedure for solving a problem.

Suppose our problem is that we want to prepare fried eggs. The algorithm could be defined as follows: Heat a pan with oil. Fry two eggs in it for seven minutes. Season with salt and pepper. In this case, the algorithm is a set of cooking instructions, the recipe if you will.

Nowadays, we tend to associate the word 'algorithm' with a computer code that provides a set of instructions for a software application. It is important to note that each step of an algorithm is precisely defined and leaves no room for uncertainty. This can be seen in the popular 3-dimensional puzzle, the "Rubik's Cube".



Solving the Rubik's Cube does not require any complex mathematical knowledge or strong spatial thinking – you only need to know when to apply which algorithm.

The reason is that each coloured square of the Rubik's Cube has a fixed number of possible moves. If you repeat a certain move often enough, the square will end up where it started.

In the Rubik's Cube, the algorithm describes how often and in which direction a certain side of the cube must be turned so that a certain square or combination of squares reaches the desired position.

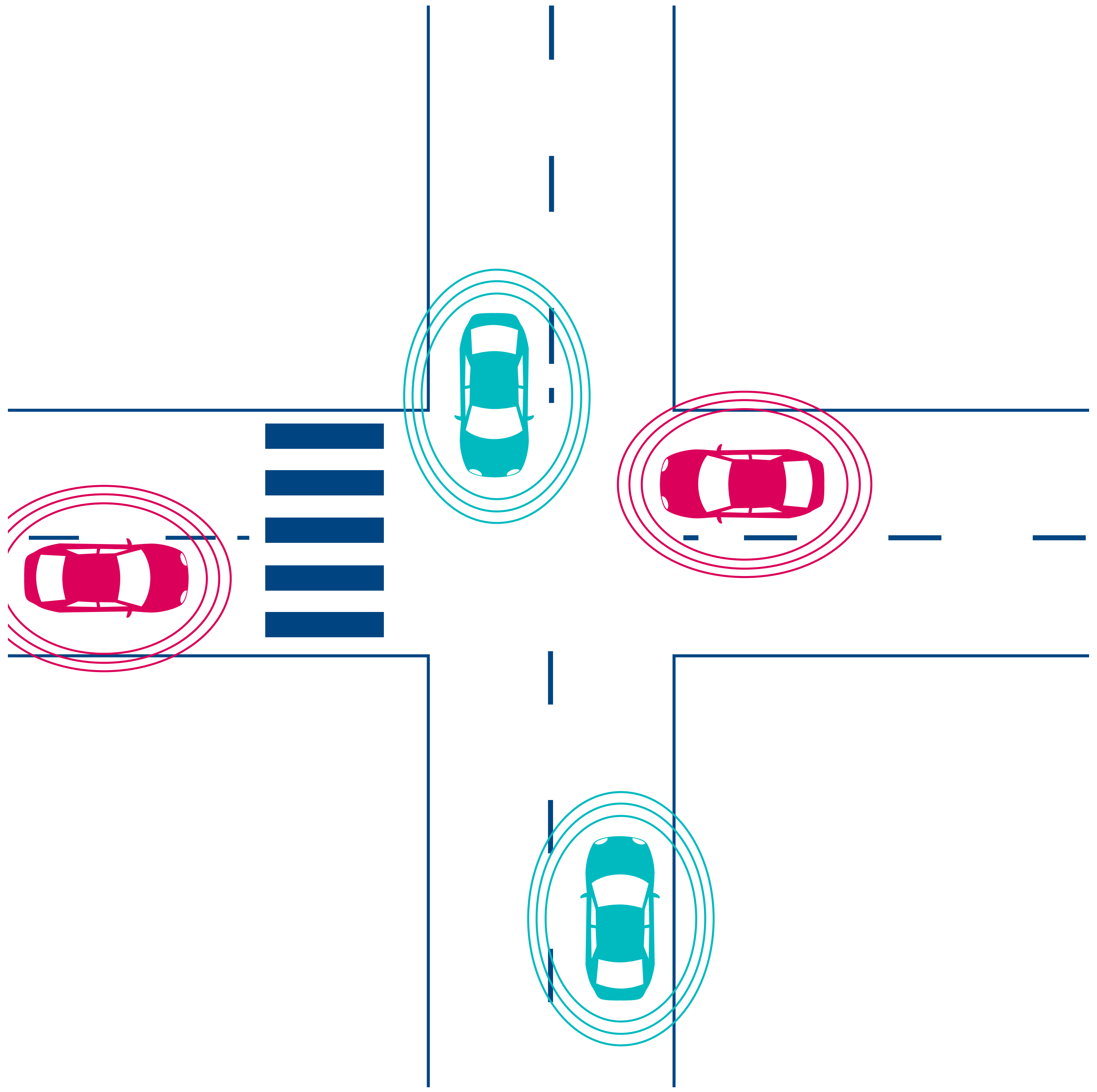
In our graphical example, it is already sufficient to turn the right side down once – a very simple algorithm.

Did you know?
Autonomous and automated

Autonomous

AI works autonomously when it has to develop an individual solution in a situation with dynamically changing environmental conditions. This situation arises when only a limited set of causal relationships can be programmed into the algorithm. One example is the case of autonomous driving.

An autonomous AI system (the self-driving car), must be able to react to other road users. While pedestrians and drivers may observe certain traffic rules, their exact actions and positions are unpredictable and are therefore difficult to program into the algorithm. In this case, an AI will act autonomously.



Automated

Accordingly, an AI works in an automated manner when all occurring causal relationships for solving the problem have been pre-programmed. This includes applications to chain different activities together, for example a repetitive daily routine.

In the graphic, the red car brakes in front of a crosswalk, regardless of whether someone or something is crossing it. Braking automatically occurs when the appropriate traffic sign is encountered.

The second braking process is to prevent a collision with the turquoise car. This process is autonomous.

› Back
› Menu

Great, you have successfully clicked to a new slide via a mouseover. Click "Back" to return to the introduction.

› Back

By the way, if you're currently using a smartphone or tablet, you can swipe and touch your way through the slides.

KONRAD ADENAUER STIFTUNG

Imprint

Publisher:
Konrad-Adenauer-Stiftung e. V. 2021, Berlin

Responsible at the Konrad Adenauer Foundation:

Jason Chumtong
Policy Advisor Artificial Intelligence
T +49 30 / 26996-3989
jason.chumtong@kas.de

#KAS4 INNOVATION

This publication of the Konrad-Adenauer-Stiftung e. V. is for information purposes only. It may not be used by political parties or by election canvassers or campaigners for the purpose of election canvassing. This applies to federal, state and local elections as well as elections to the European Parliament.

Illustrations, animations, design and typesetting: Franziska Feldmann Grafikdesign | Webdesign

Produced with the financial support of the Federal Republic of Germany.

ISBN: 978-3-95721-973-2

› Back
› Menu