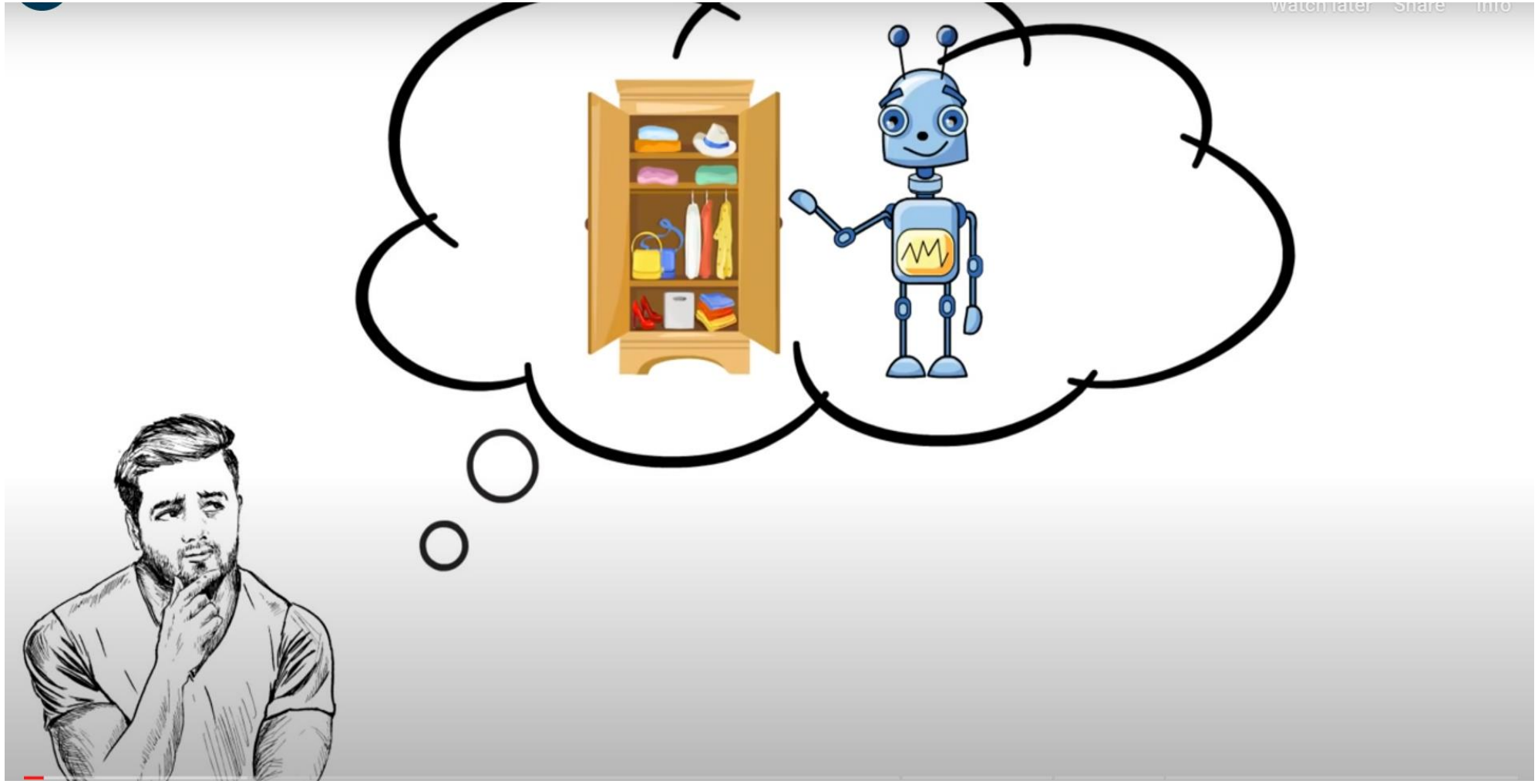
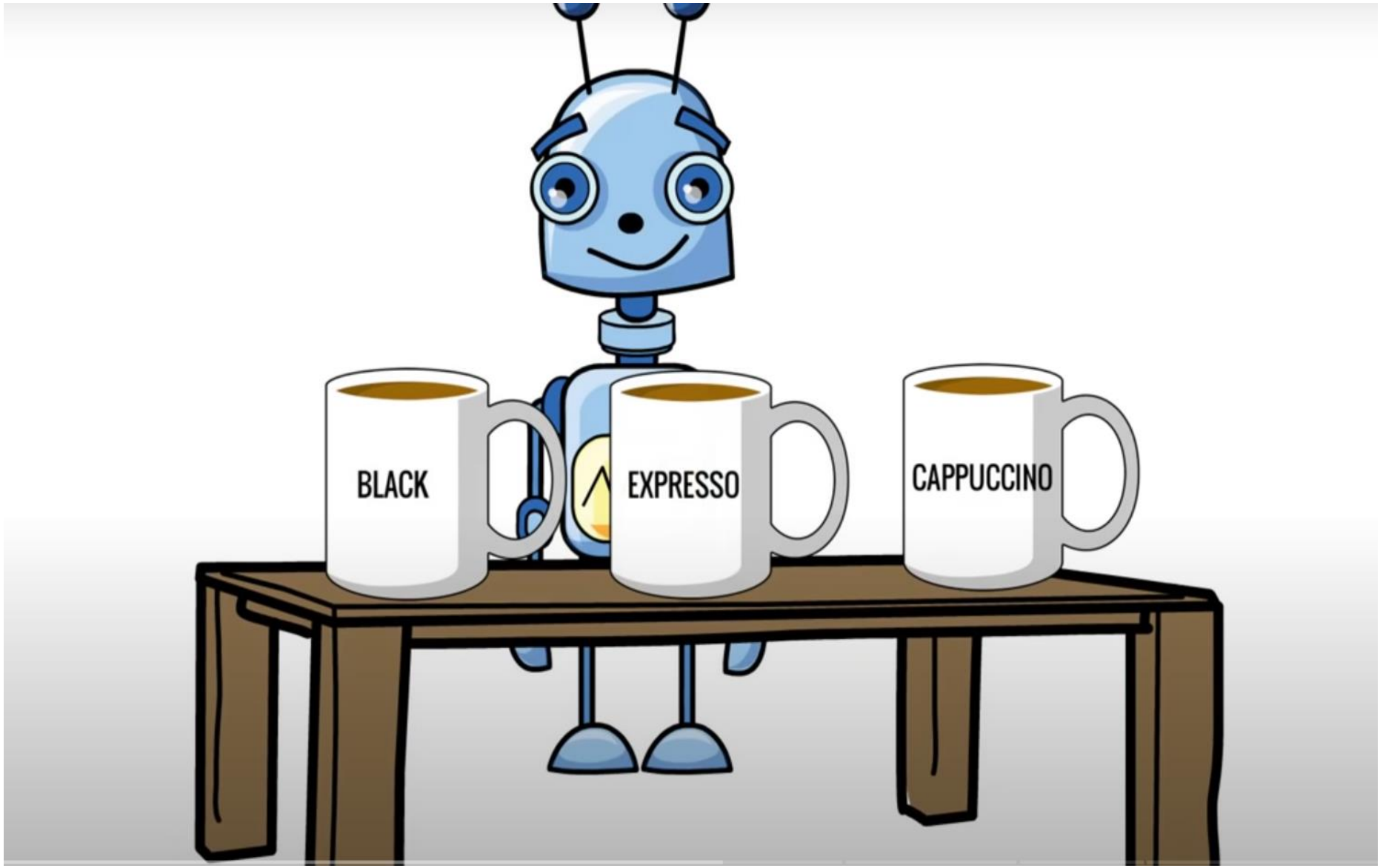
Introduction to AI Ethics

Contents:

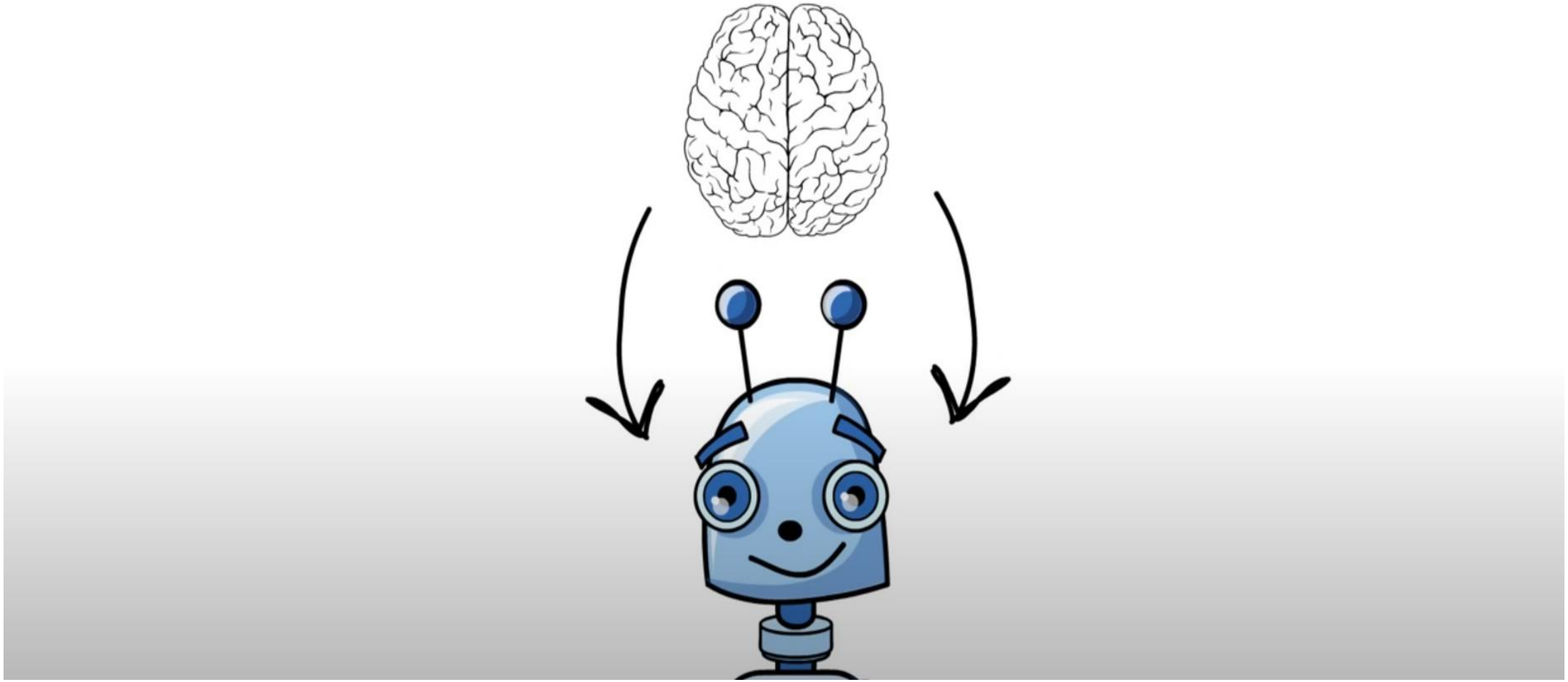
- Artificial intelligence
- Ways of implementing AI
- Advantages and disadvantages of AI
- Definition of morality and ethics
- Descriptive ethics
- Normative ethics
- Meta-ethics
- Applied ethics
- Impact on society, human psychology, legal system, Environment and planet, trust
- Challenges of AI data governance
- Ethical implications and responsibilities.

Artificial Intelligence:

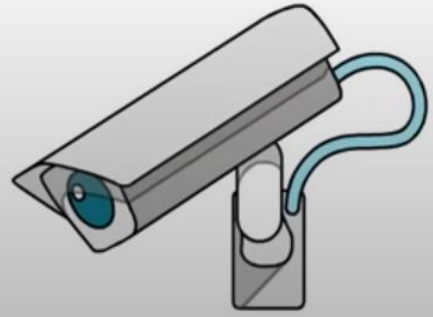



BLACK
EXPRESSO
CAPPUCCINO

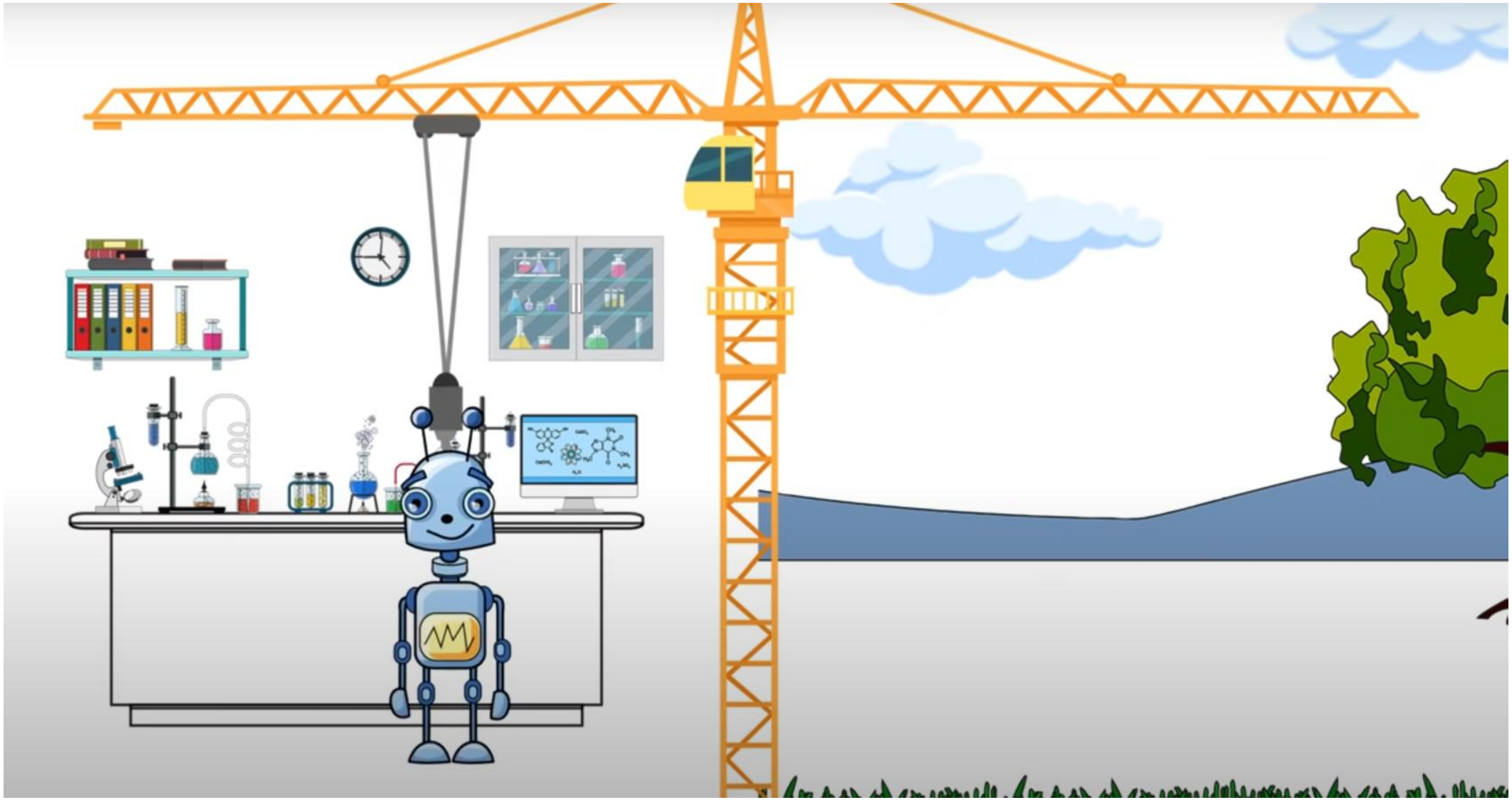
THESE ARE THE PRODUCTS OF ARTIFICIAL INTELLIGENCE

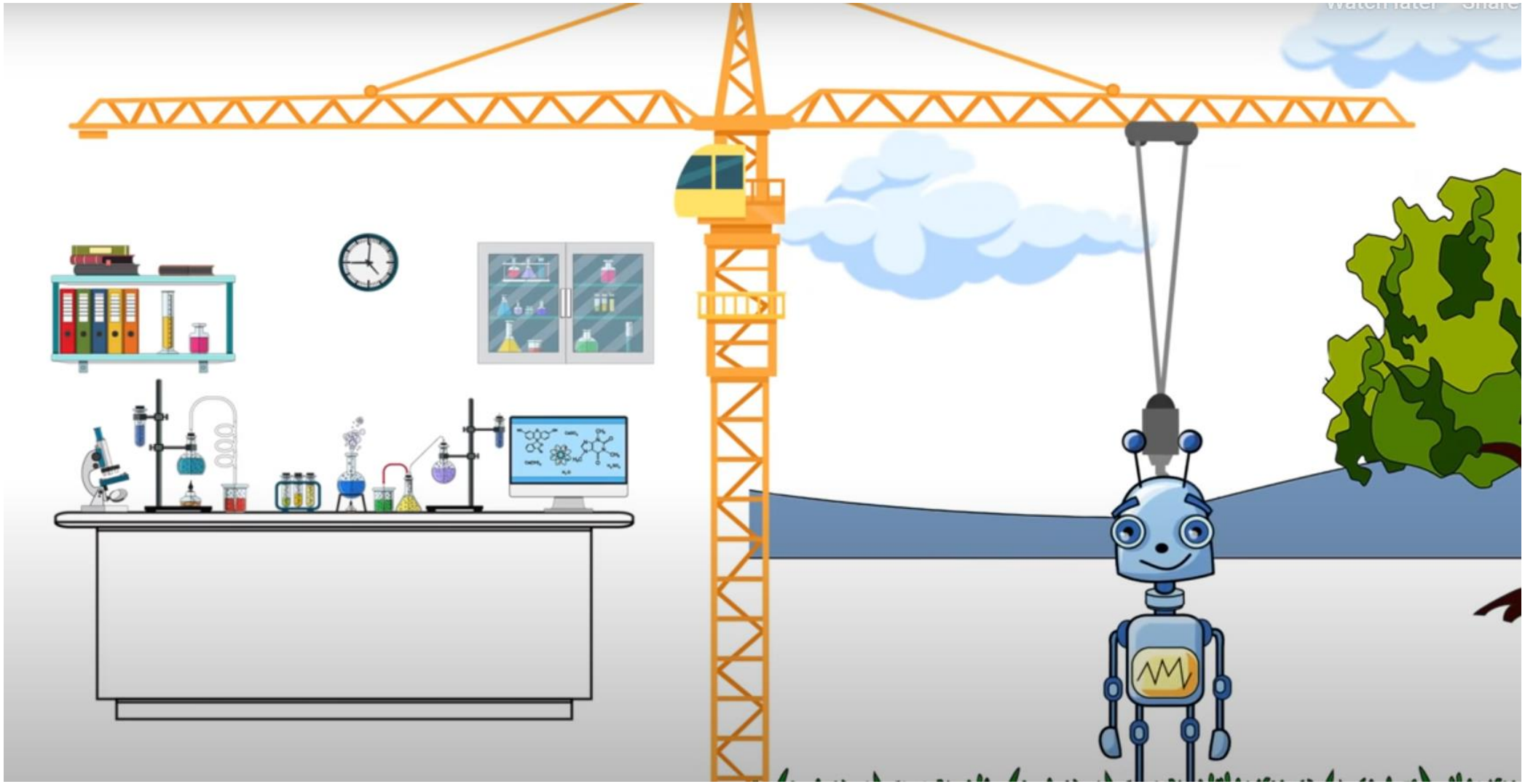


AI MAY NOT BE AS OBVIOUS AS IN THE PREVIOUS EXAMPLES

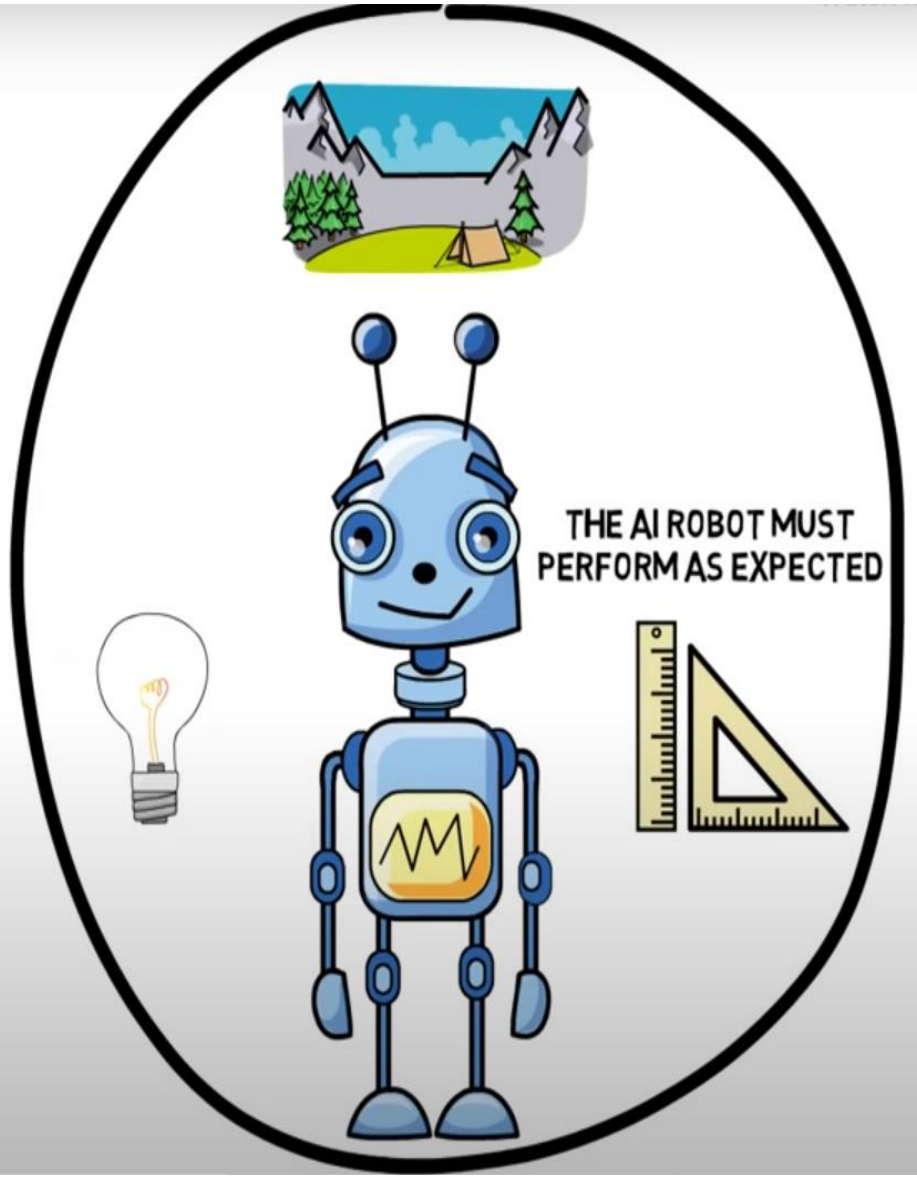


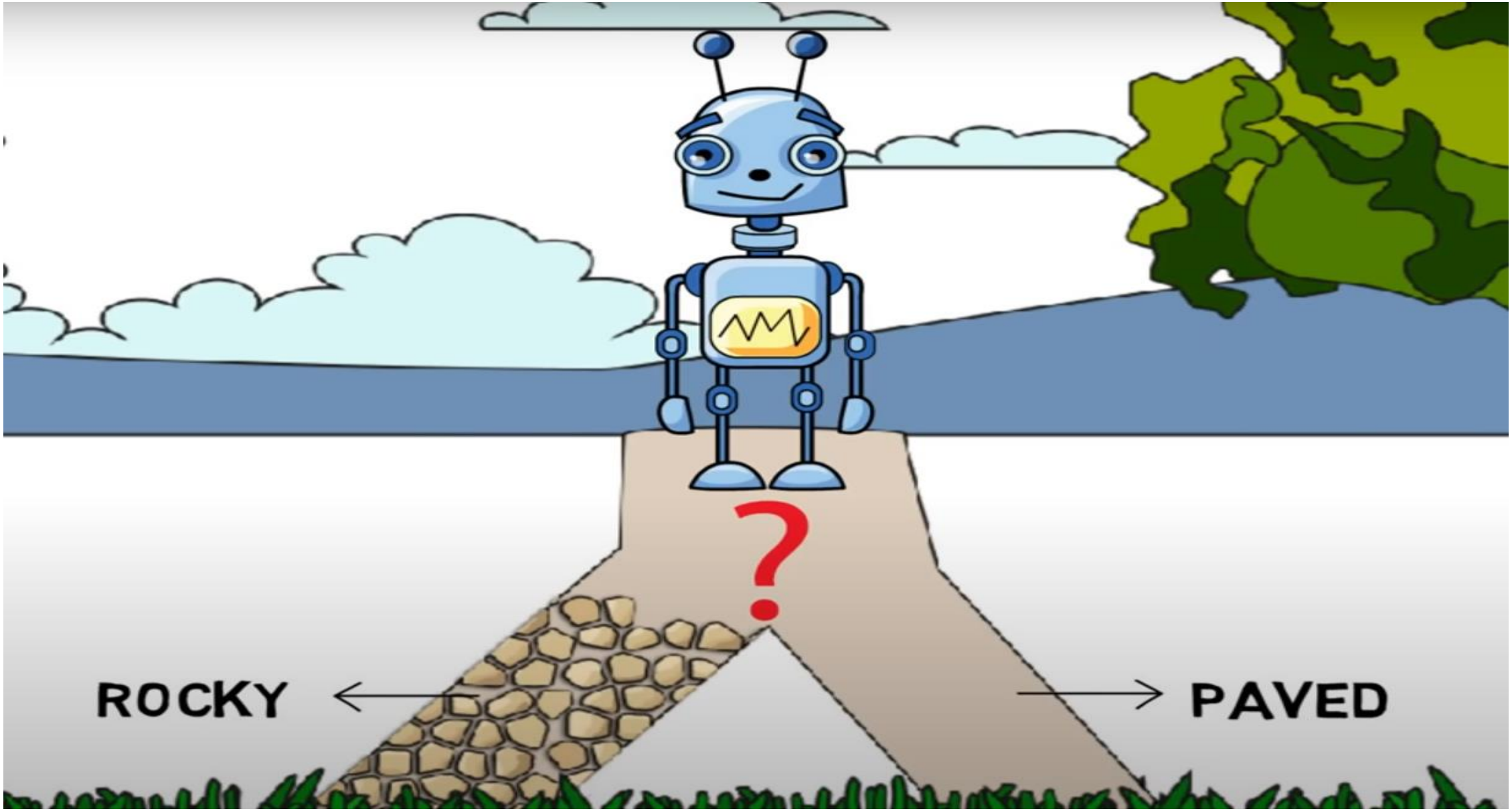
WHAT DOES AN AI DO AT ITS CORE?



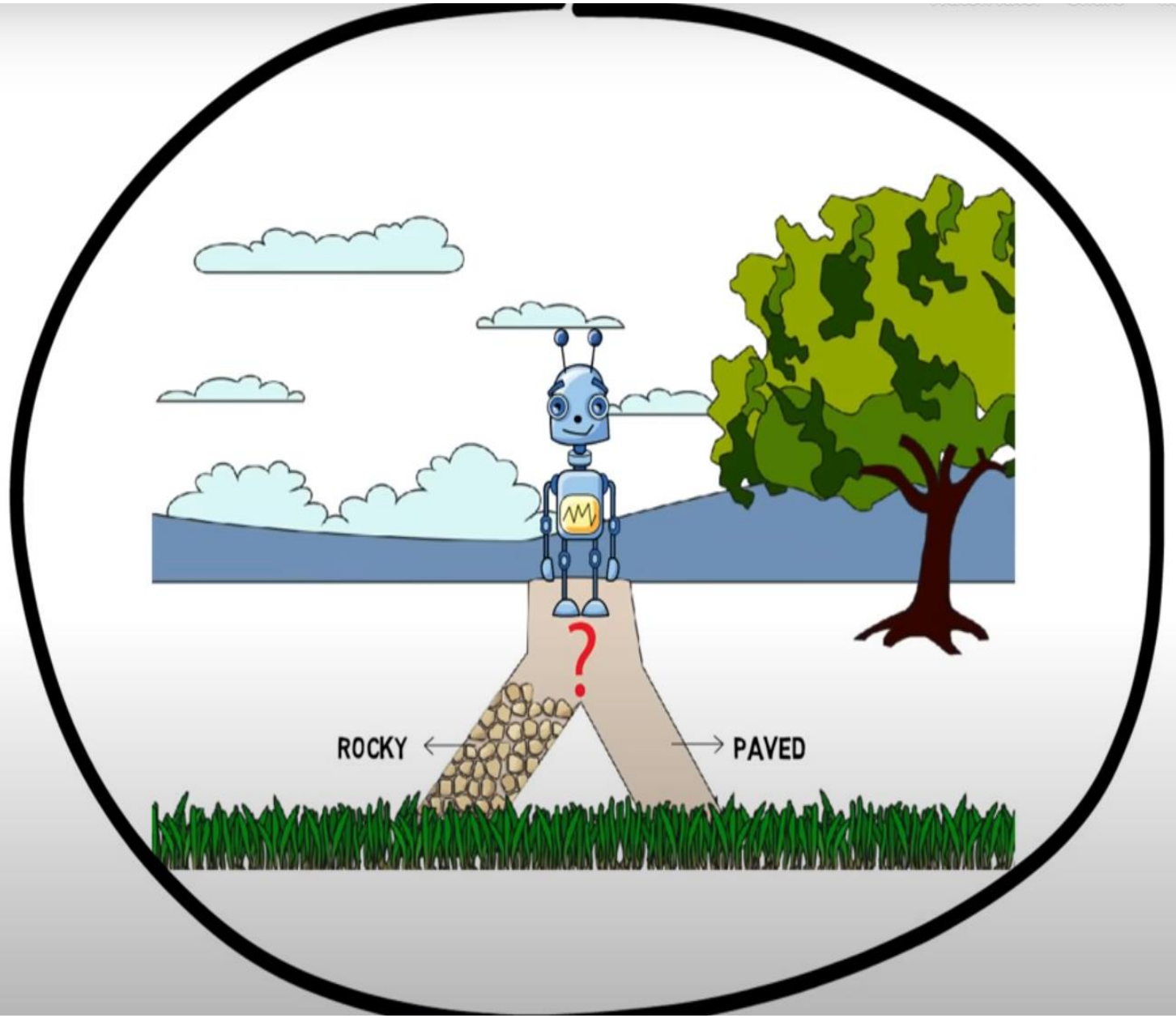


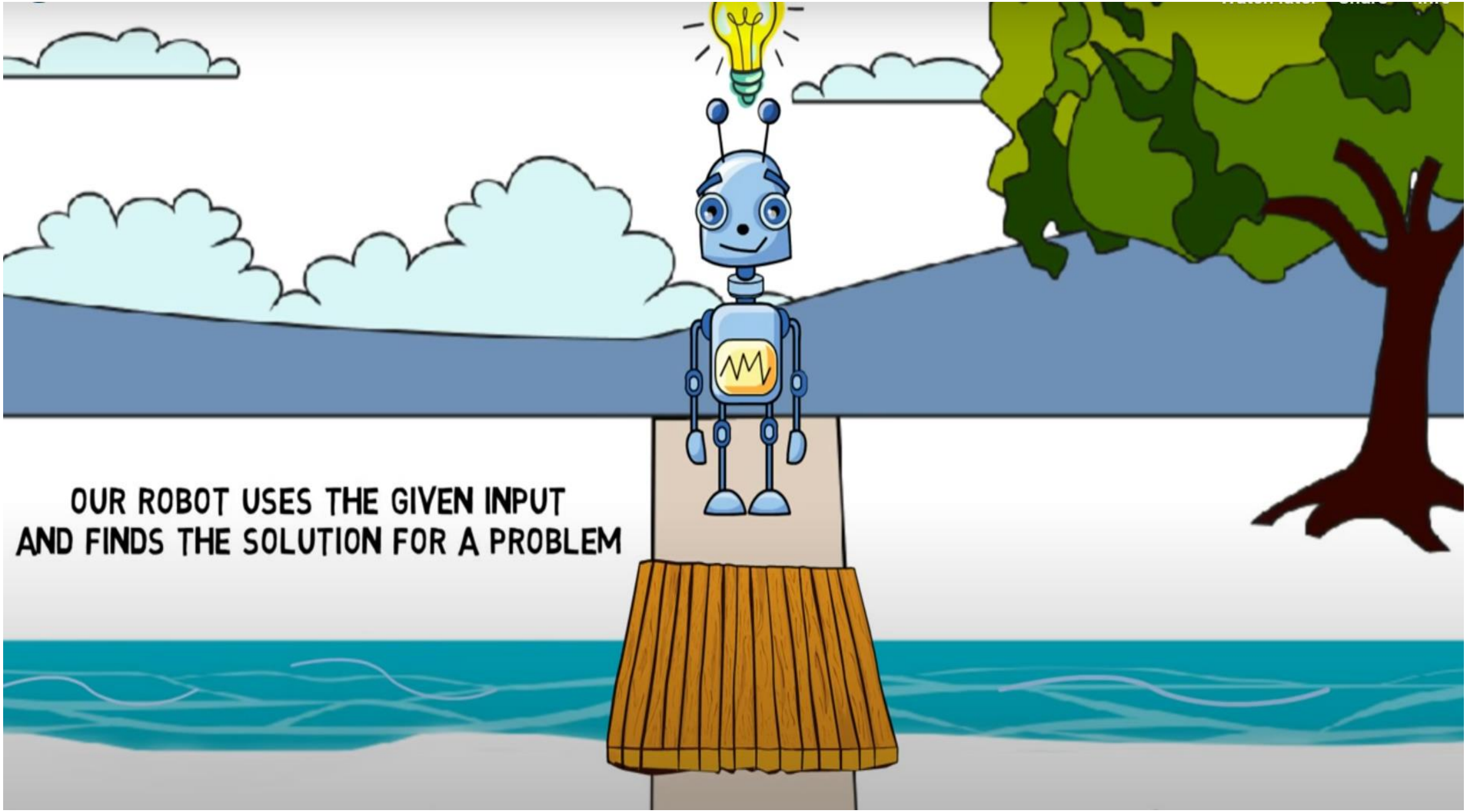
GENERALIZED LEARNING



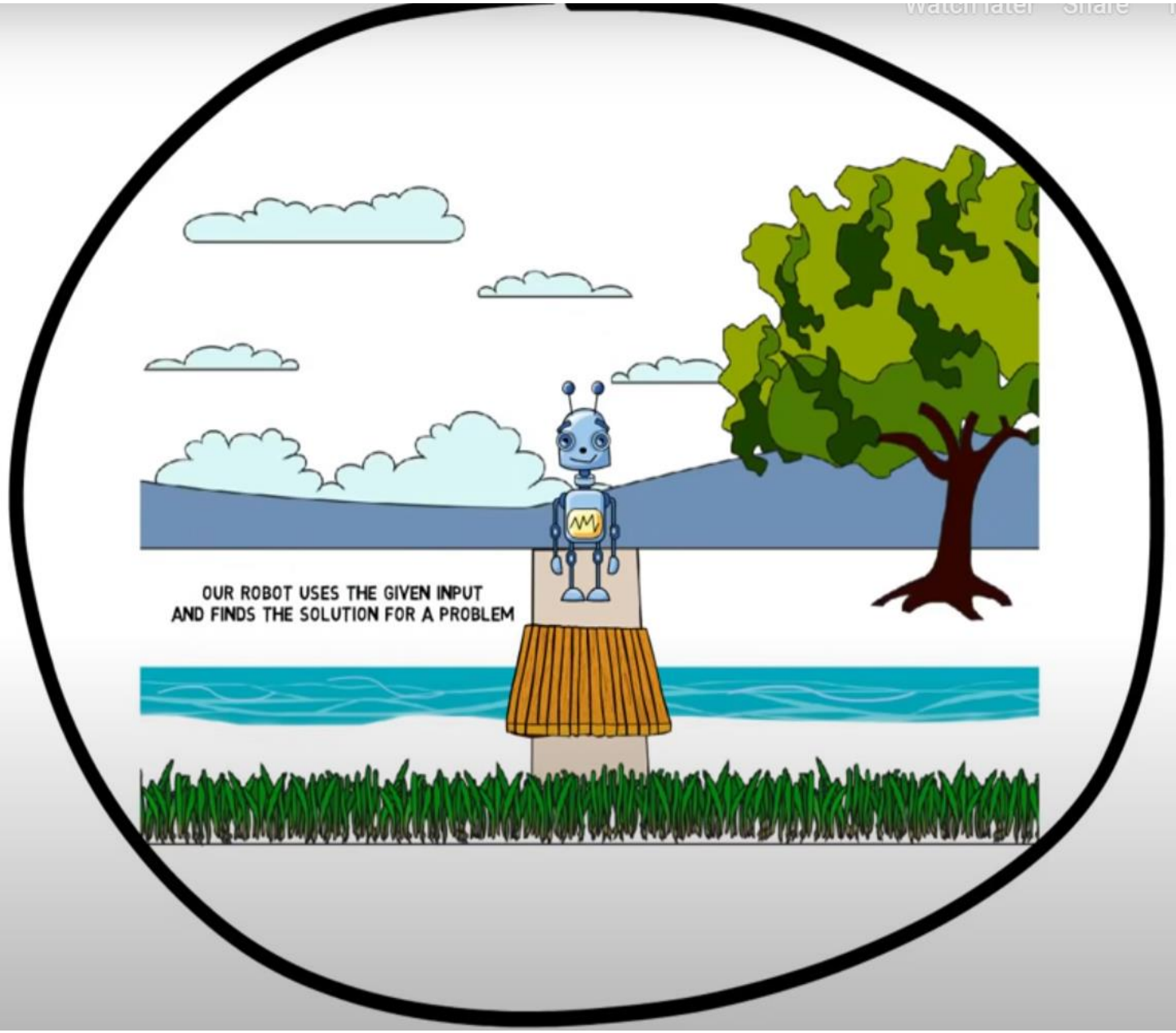

?
ROCKY
PAVED

REASONING



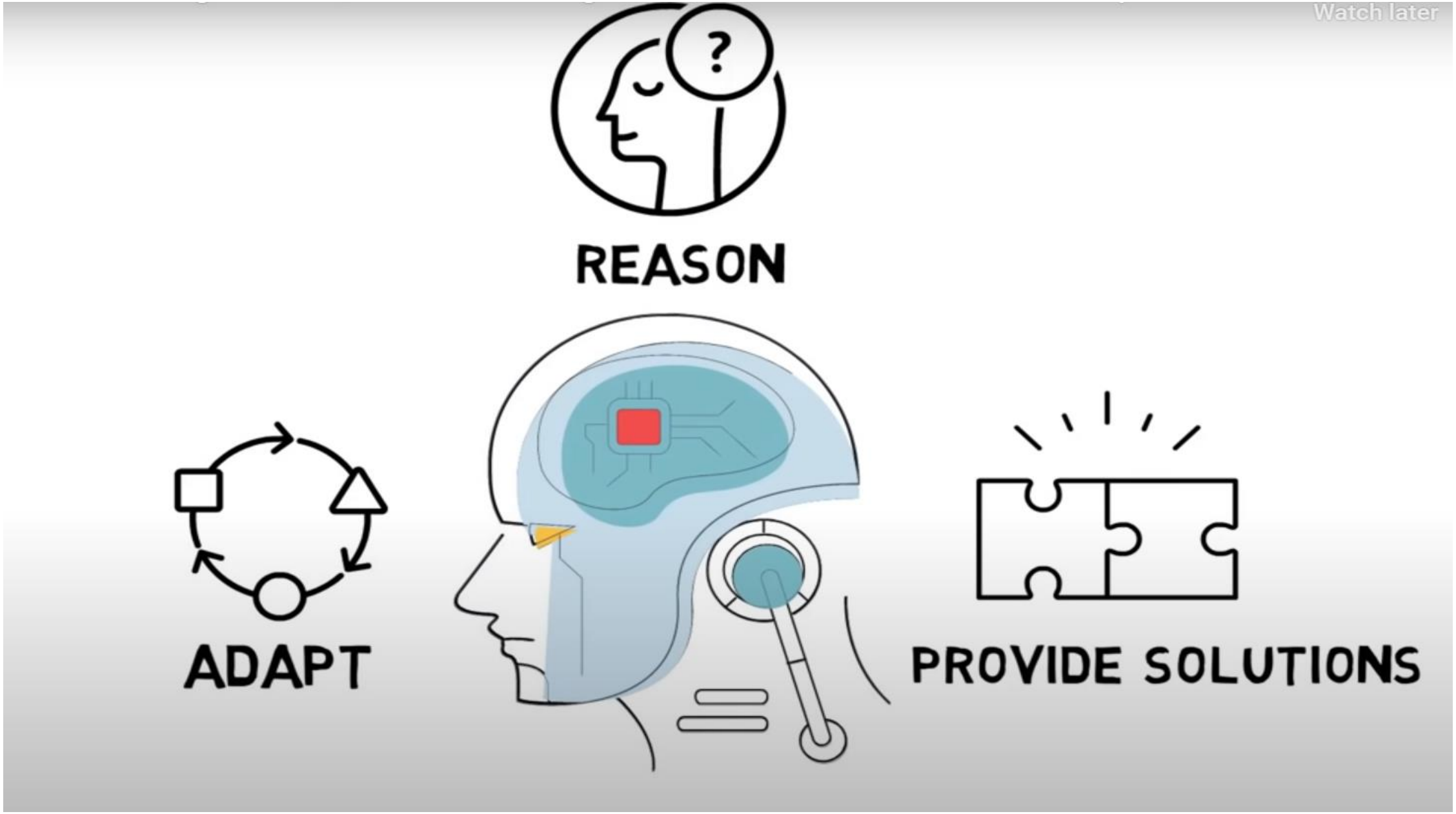

OUR ROBOT USES THE GIVEN INPUT
AND FINDS THE SOLUTION FOR A PROBLEM

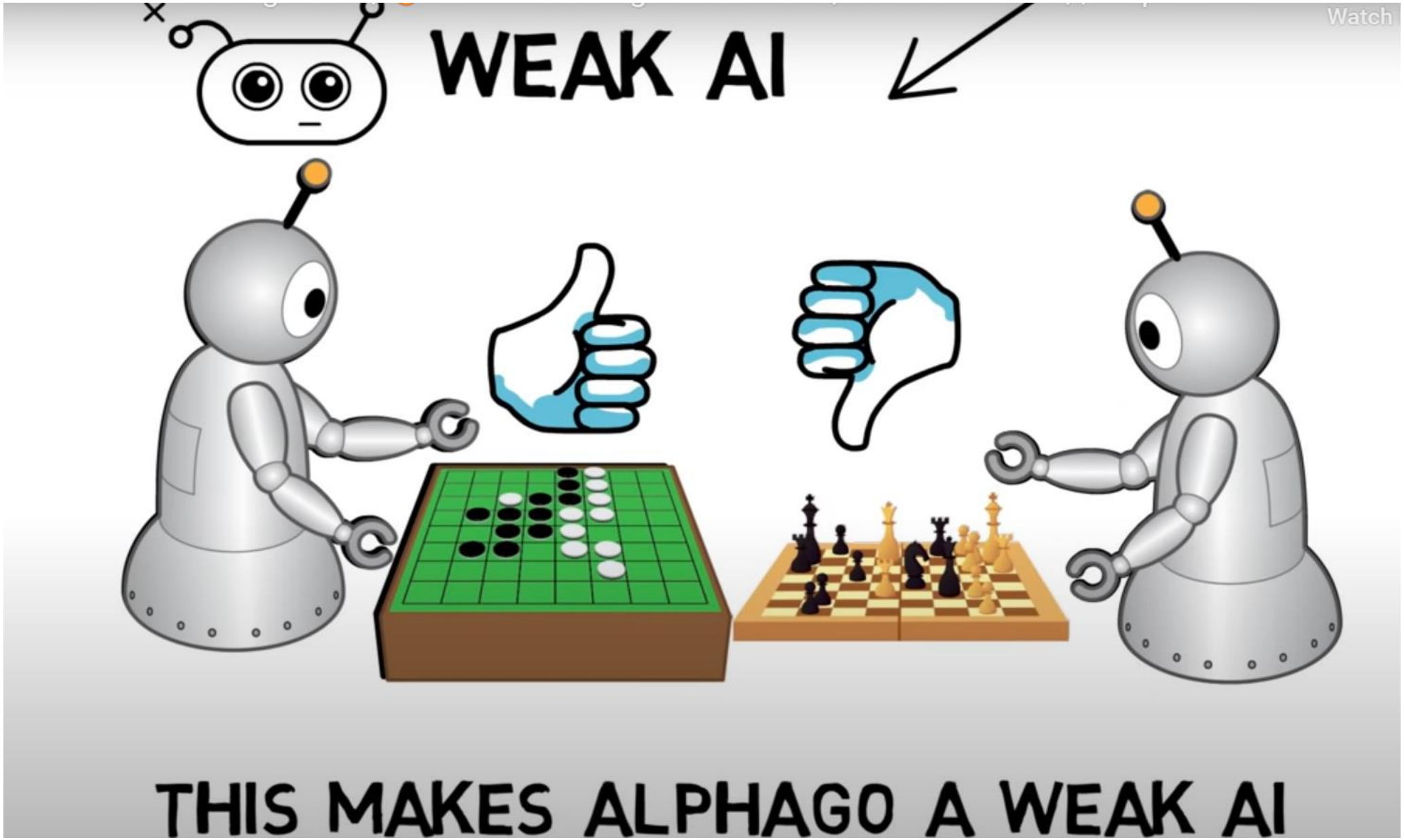
PROBLEM SOLVING



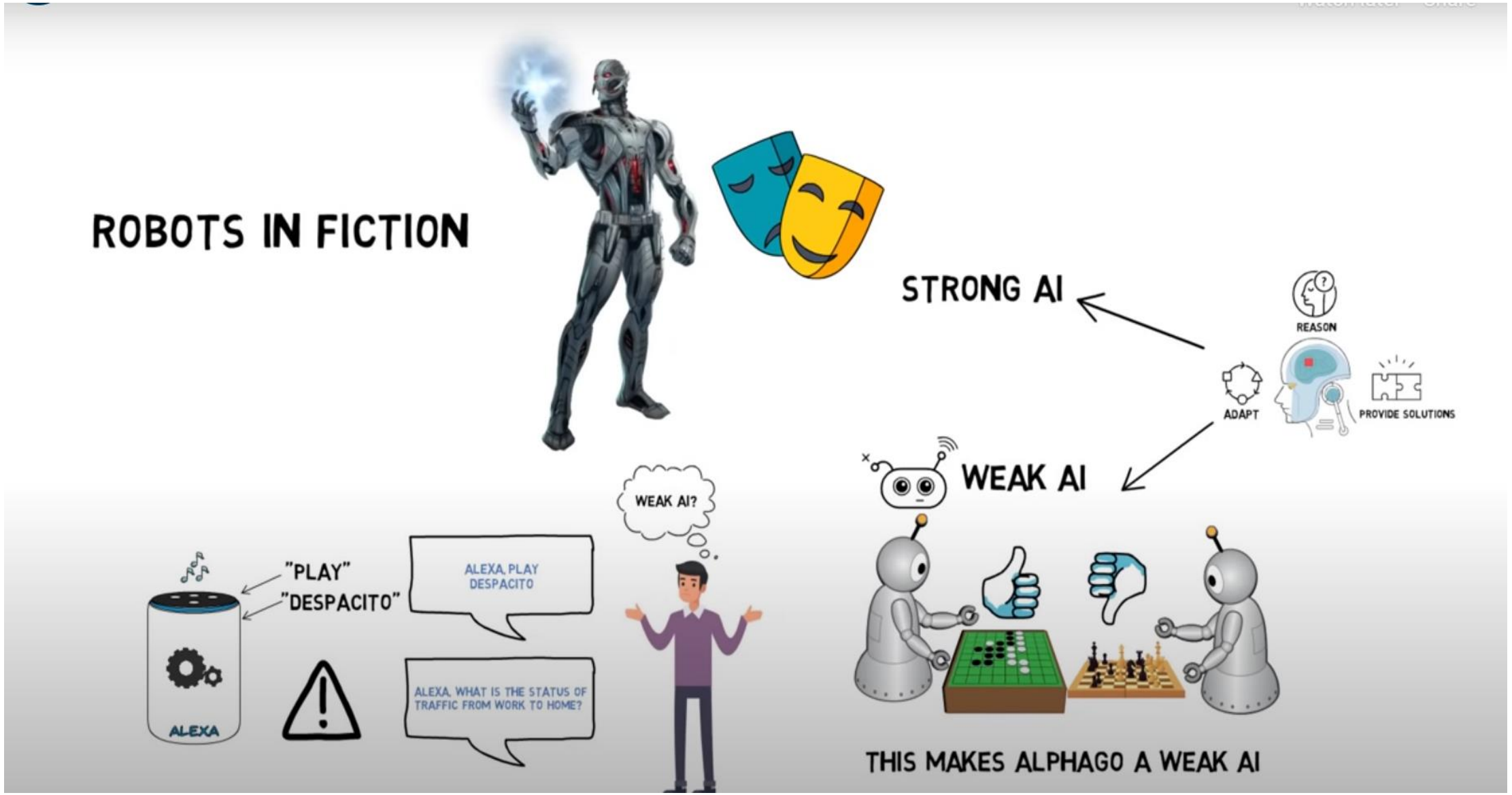
- [x] Generalized learning
- [x] Reasoning
- [x] Problem solving

ARTIFICIALLY INTELLIGENT


Watch later
REASON
ADAPT
PROVIDE SOLUTIONS


WEAK AI
THIS MAKES ALPHAGO A WEAK AI

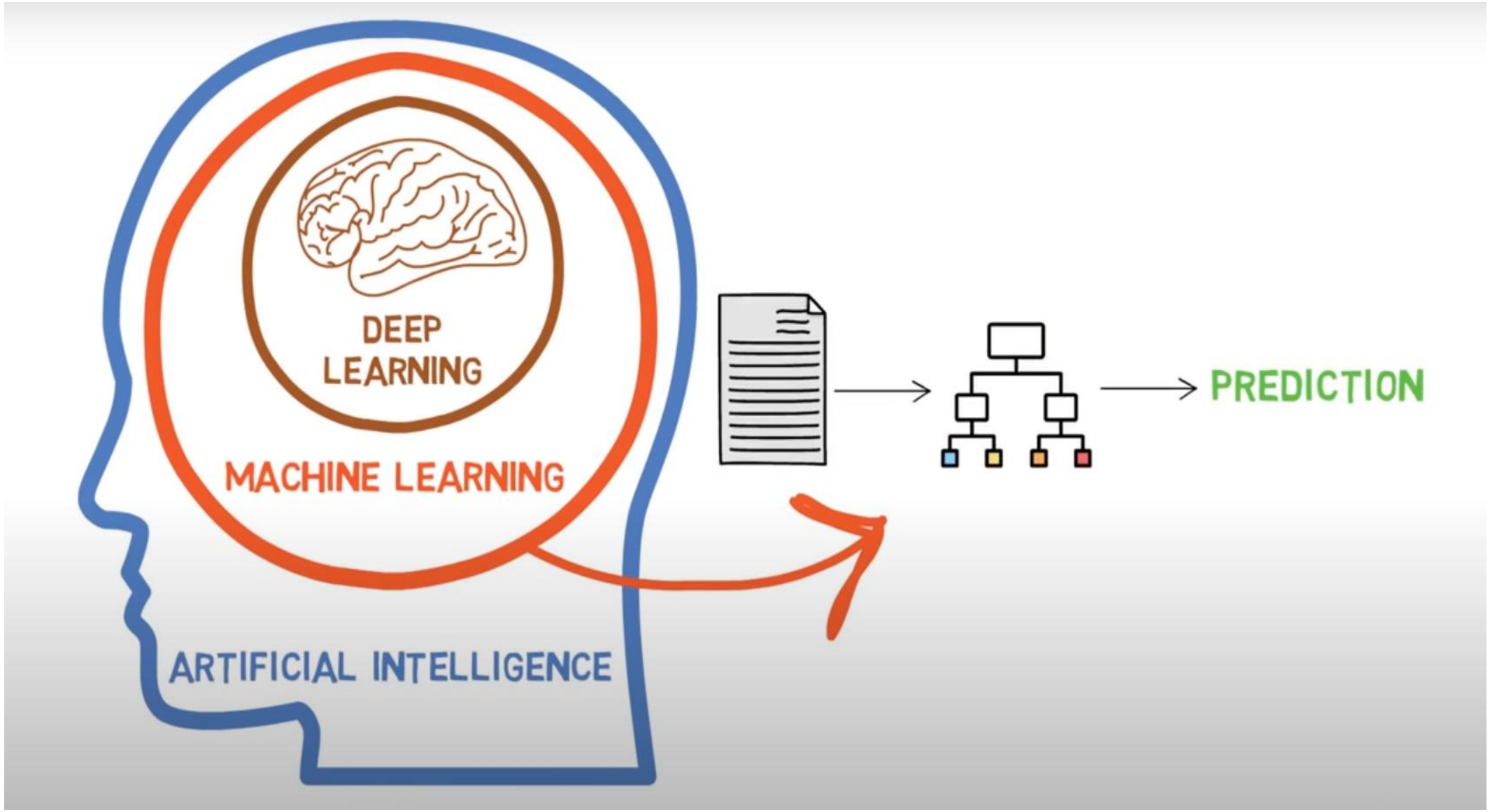

"PLAY"
"DESPACITO"
ALEXA
ALEXA, PLAY DESPACITO
ALEXA, WHAT IS THE STATUS OF TRAFFIC FROM WORK TO HOME?
WEAK AI?

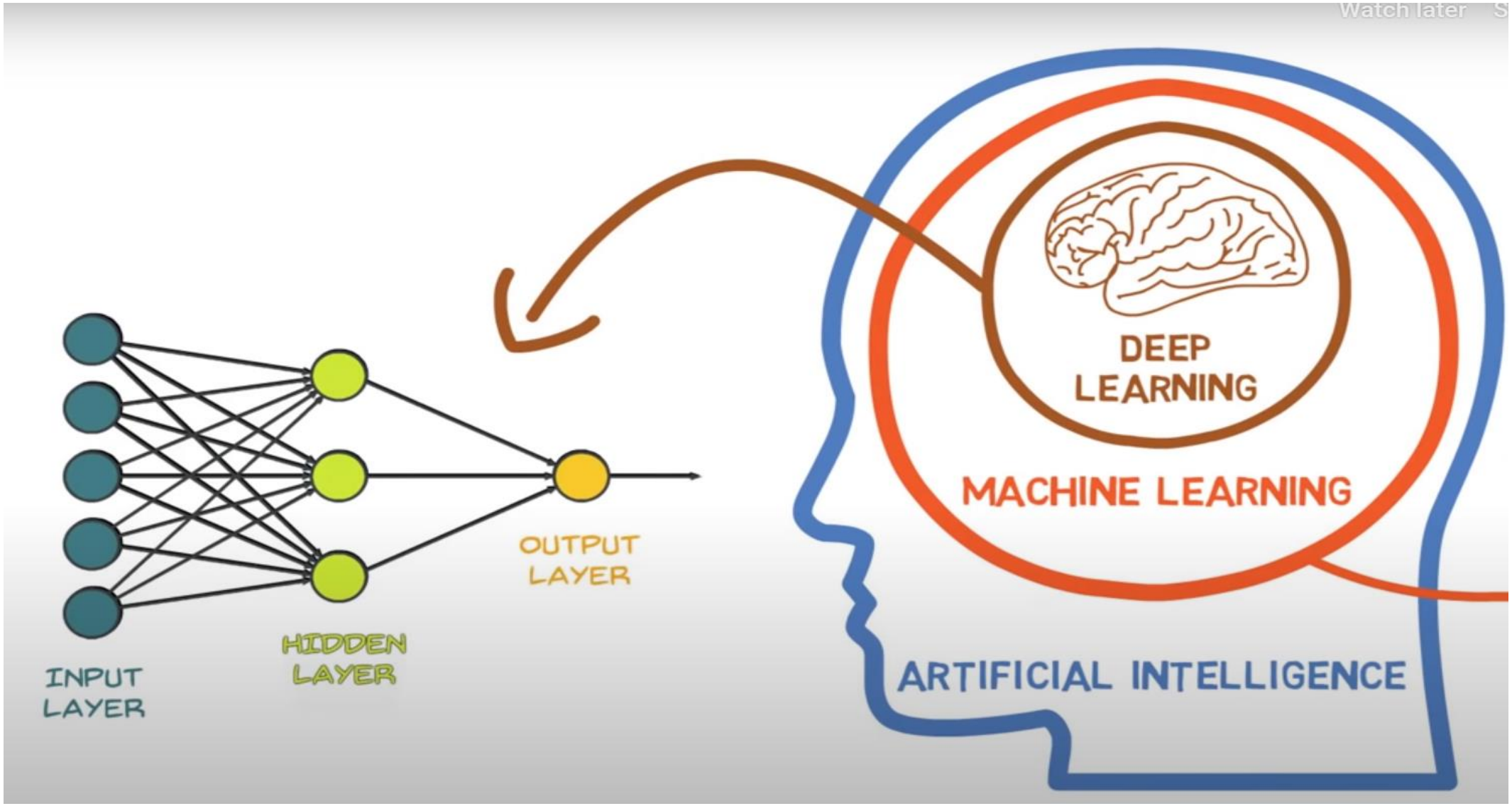

ROBOTS IN FICTION
STRONG AI
REASON
ADAPT
PROVIDE SOLUTIONS
WEAK AI
"PLAY"
"DESPACITO"
ALEXA
ALEXA, PLAY DESPACITO
ALEXA, WHAT IS THE STATUS OF TRAFFIC FROM WORK TO HOME?
WEAK AI?
THIS MAKES ALPHAGO A WEAK AI

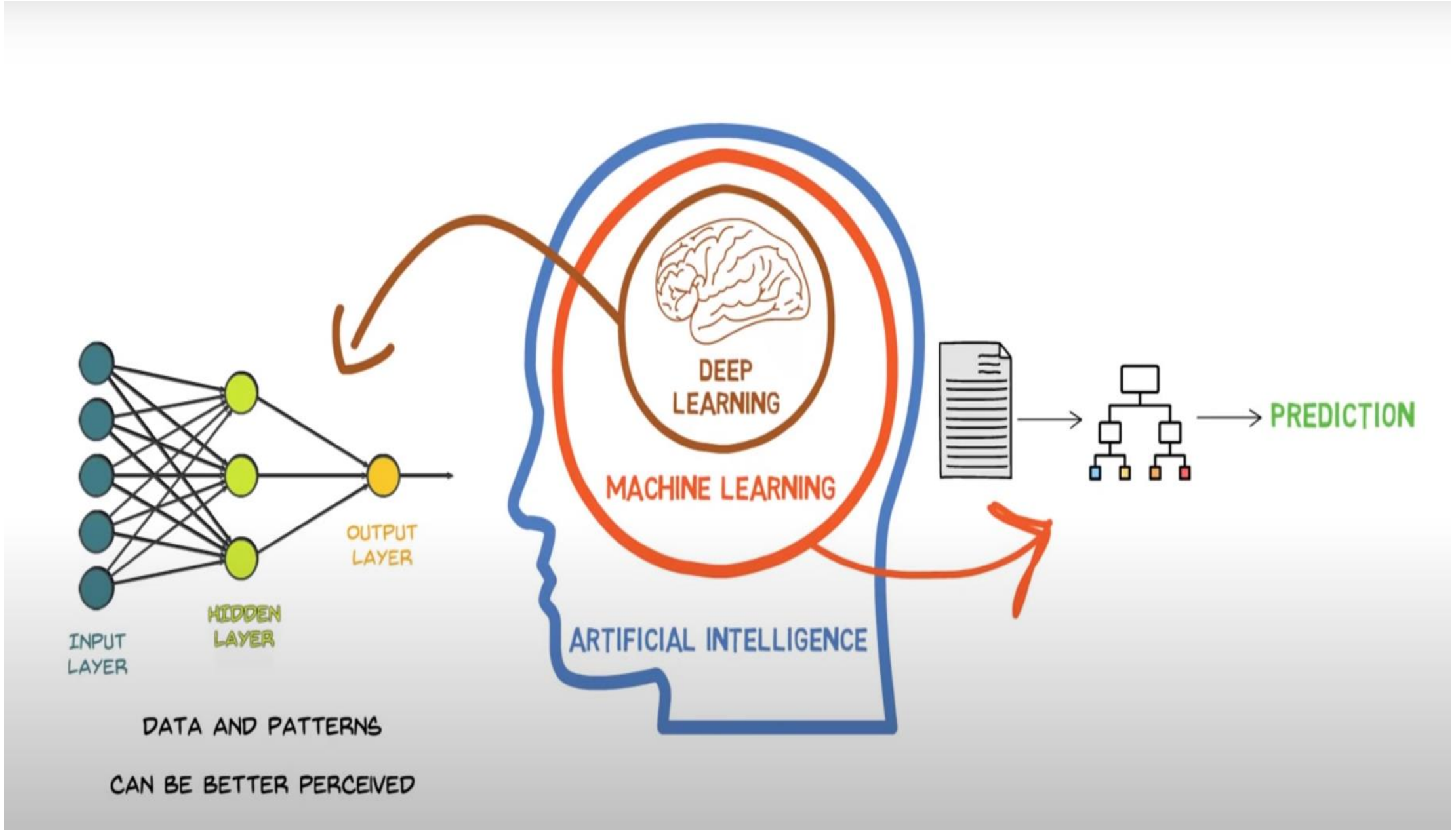
MACHINE LEARNING

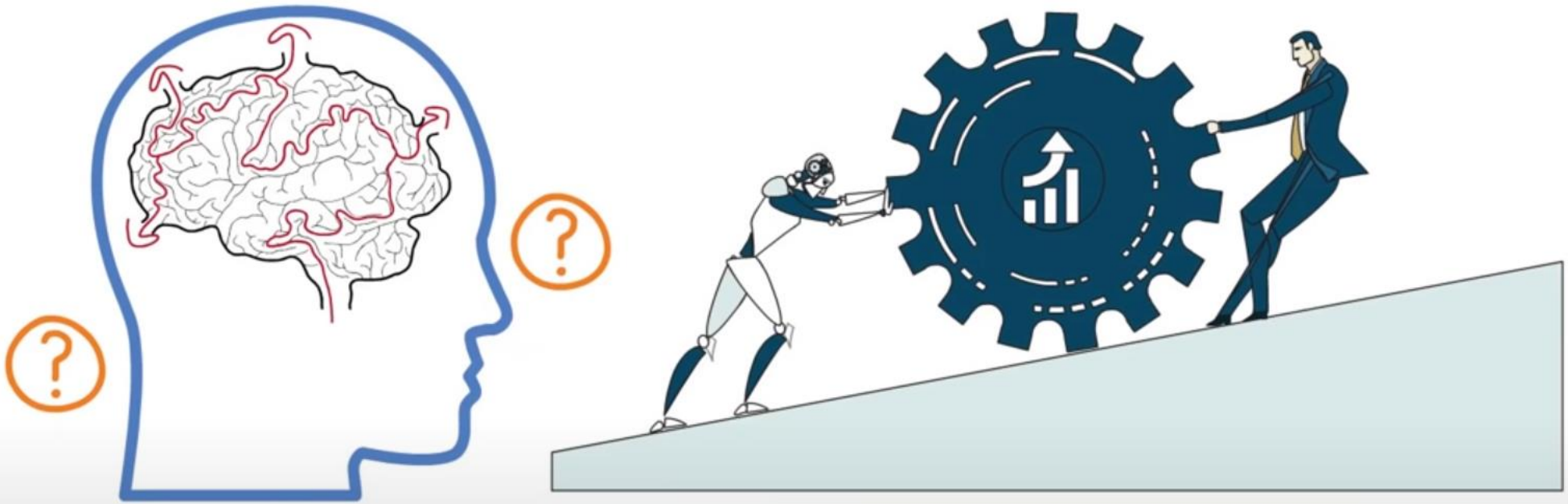
ARTIFICIAL INTELLIGENCE

DEEP LEARNING


DEEP
LEARNING
MACHINE LEARNING
ARTIFICIAL INTELLIGENCE
PREDICTION


Watch later
INPUT LAYER
HIDDEN LAYER
OUTPUT LAYER
DEEP LEARNING
MACHINE LEARNING
ARTIFICIAL INTELLIGENCE


DEEP
LEARNING
MACHINE LEARNING
ARTIFICIAL INTELLIGENCE
PREDICTION
INPUT
LAYER
HIDDEN
LAYER
OUTPUT
LAYER
DATA AND PATTERNS
CAN BE BETTER PERCEIVED



HOWEVER, WITH THE MATURATION OF TECHNOLOGY, WE CAN ONLY WAIT AND WATCH WHAT THE FUTURE OF AI HOLD FOR US

Ways of implementing AI:

1. Data Acquisition
   - Data collection
   - Data Pipeline
   - Data validation, clean up
   - EDA
2. Model Development
   - Feature engineering
   - Model training
   - Model Evaluation

Ways of implementing AI:

3. Deployment
    - Product Integration
    - Validation
4. Monitoring

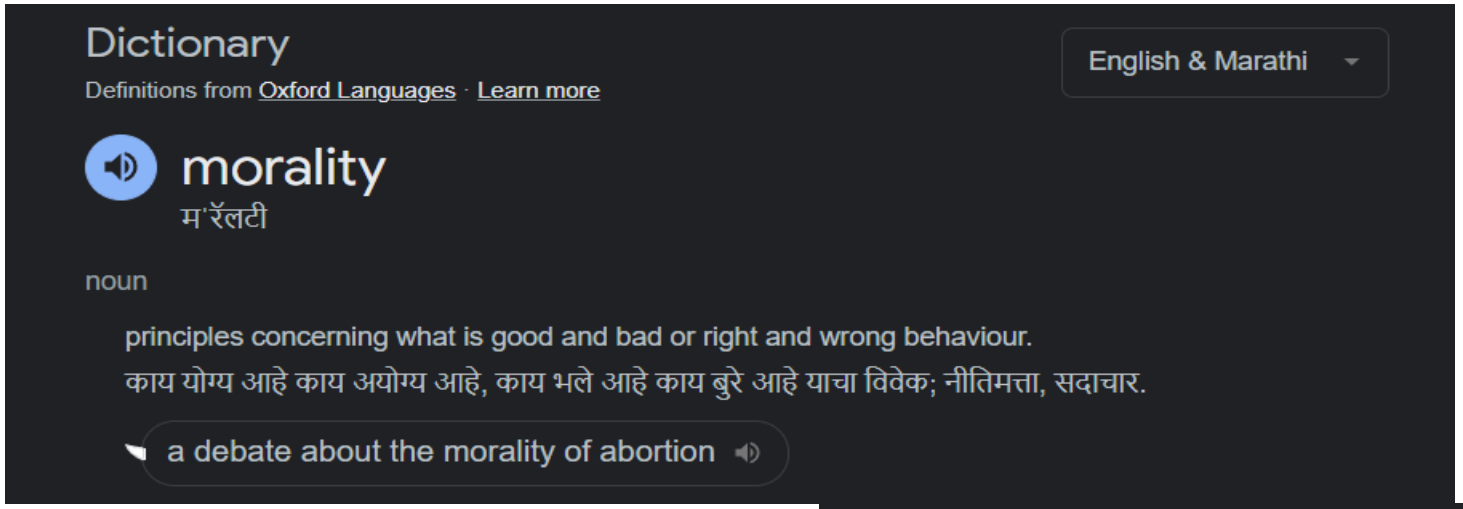
Advantages of AI:

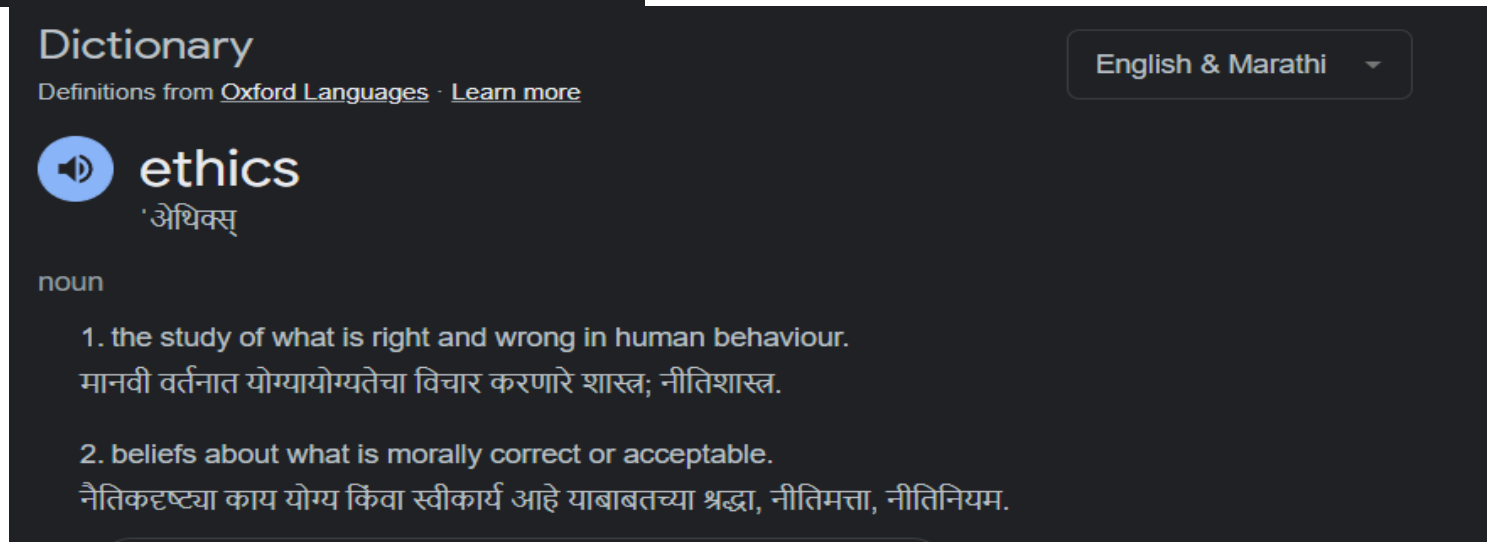
1. Reduction in human error
2. Zero Risks
3. 24*7 Availability
4. Digital Assistance
5. New Inventions
6. Unbiased Decisions
7. Perform Repetitive Jobs
8. Daily Applications

Disadvantages of AI:

1. High cost
2. No creativity
3. Unemployment
4. Make Humans Lazy
5. No Ethics
6. Emotionless

Definition of morality and Ethics:






Sancity
MORAL
FOUNDATIONS
care
Loyalty
respect
fairness


integrity
principles
moral
value
honesty
honor
ETHICS
right
choice
fairness
conscience
responsibility

ETHICS VS MORALS

In most cases, when any one of us does something, we try to act according to what society believes is right. More often, we listen closely to what our own beliefs about right or wrong are telling us, even if they're different from society's views. These two have to do with ethics and morals.

MEANING

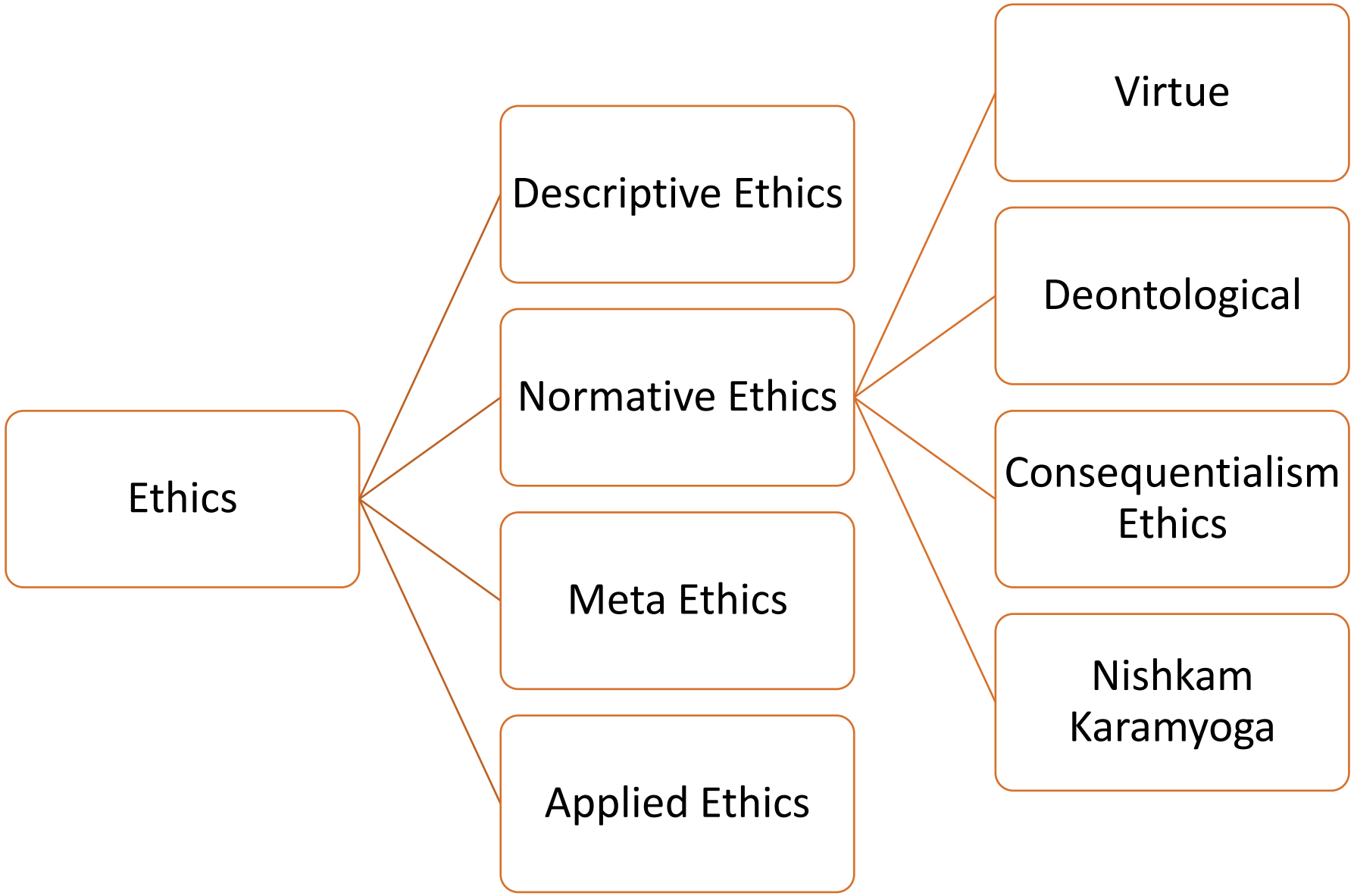
ETHICS refers to the rules that a social system provides us with.

MEANING

MORALS are our own principles.

Ethics:

1. Ethics are based on our understanding of what is right and wrong
2. Ethics involves making choices about how we should act
3. Ethics require us to think about the consequences of our actions
4. Ethics often involve trade-offs between competing values or interests


Ethics
Descriptive Ethics
Normative Ethics
Meta Ethics
Applied Ethics
Virtue
Deontological
Consequentialism Ethics
Nishkam Karamyoga

1. Descriptive Ethics:

- It deals with what ***people actually believe (or are made to believe) to be right or wrong.***
- Accordingly holds up the human actions acceptable or not acceptable or punishable under custom law.
- **Descriptive ethics** is also called **Comparative ethics** because it compares the ethics of the past and present.
- It also takes inputs from other disciplines such as **anthropology, psychology, sociology**, and **history** to explain moral right or wrong.

2.Normative ethics:

- Normative Ethics deals with “norms” or a set of considerations about how one should act.
- It is a *study of “ethical action”* and *sets out the rightness or wrongness* of the actions.
- It is also called **prescriptive ethics.**
- The Golden rule of normative ethics is “doing to others as we want them to do to us”
- Normative ethics also provides justification for punishing a person who disturbs social and moral order.

2.Normative ethics:

1. **Virtue Ethics:**

   Virtue ethics focuses on one's character and the virtues for determining or evaluating ethical behavior.

   Plato, Aristotle, and Thomas Aquinas were major advocates of Virtue ethics. Plato gave a scheme of *four cardinal virtues viz. prudence, justice, temperance and fortitude (courage*).

2. **Deontological ethics:**

   Deontology is an ethical theory that says actions are good or bad according to a clear set of rules. Its name comes from the Greek word deon, meaning duty. Actions that align with these rules are ethical, while actions that don't aren't.

2.Normative ethics:

3. Consequentialism:

Consequentialism is a theory that says whether something is good or bad depends on its outcomes. An action that brings about more benefit than harm is good, while an action that causes more harm than benefit is not. The most famous version of this theory is utilitarianism.

4. Nishkam Karamyog:

An important philosophical concept, in karma yoga it means to act unselfishly, or without personal game in the mind. When acting out in Nishkam Karma, an individual is acting without any expectation that good will return to him/her.

Meta-Ethics:

- Meta Ethics or "*analytical ethics*" deals with the *origin of the ethical concepts themselves.*
- Which studies the very foundation of morality itself.
- Meta ethics ask questions as basic as what is morality? What's it's nature.
- It is basically a *highly abstract way of thinking about ethics*.

Applied Ethics:

- *Decision ethics*
- *Professional ethics*
- *Clinical Ethics*
- *Business Ethics*
- *Organizational ethics*
- *social ethics*

Impact on society:

1. The labour market
2. Inequality
3. Privacy human rights and dignity
4. Democracy

1. **The labour Market**

   - People have been concerned about the displacement of workers by technology for centuries.

   - One study asked 1,896 experts about the impact of emerging technologies; 48 percent believed that robots and digital agents would displace significant numbers of both 'blue' and 'white' collar workers.

   - However, the other half of the experts who responded to this survey (52%) expected that technology would not displace more jobs than it created by 2025 **(Smith and Anderson, 2014)**

(Brynjolfsson and McAfee, 2014):

Technological progress is going to leave behind some people, perhaps even a lot of people, as it races ahead… there's never been a better time to be a worker with special skills or the right education because these people can use technology to create and capture value.

Ford (2009)

warns that 'at some point in the future — it might be many years or decades from now — machines will be able to do the jobs of a large percentage of the 'average' people in our population, and these people will not be able to find new jobs’

(Duckworth et al., 2019)

The team surveyed 156 academic and industry experts in machine learning, robotics, and intelligent systems, and asked them what tasks they believed could currently be automated. They found that work that is clerical, repetitive, precise, and perceptual can increasingly be automated, while work that is more creative, dynamic, and human oriented tends to be less 'automatable'

2. **Inequality:**

- AI and robotics technology are expected to allow companies to streamline their businesses, making them more efficient and more productive. However, some argue that this will come at the expense of their human workforce. This will inevitably mean that revenues will be split across fewer people, increasing social inequalities.

Jack Stilgoe

-'Does AI have to increase inequality? Could you design systems that target, for example, the needs of the poorest people? If AI was being used to further benefit rich people more than it benefits poor people, which it looks likely to be, or more troublingly, put undue pressure on already particularly marginalized people, then what might we do about that? Is that an appropriate use of AI?

Nimitz

-this accumulation of power in the hands of a few — the power of money, the power over infrastructures for democracy and discourse, the power over individuals based on profiling and the dominance in AI innovation…must be seen together, and…must inform the present debate about ethics and law for AI'.

3. Privacy, human rights and dignity:

-Smith (2018), President of Microsoft, recently remarked: ‘Intelligent technology raises issues that go to the heart of fundamental human rights protections like privacy and freedom of expression. These issues heighten the responsibility for tech companies that create these products. In our view, they also call for thoughtful government regulation and for the development of norms around acceptable uses.’

- A survey of IPA customers showed that people's biggest privacy concern was their device being hacked (68.63%), followed by it collecting personal information on them (16%), listening to their conversations 24/7 (10%), recording private conversations (12%), not respecting their privacy (6%), storing their data (6%) and the 'creepy' nature of the device (4%) (Manikonda et al, 2018). However despite these concerns, people were very positive about the devices, and comfortable using them

4. Democracy:

Impact on Human Psychology

1. Relationship
2. Personhood.

Impact on Human Psychology

John Havens

-The biggest risk [of AI] that anyone faces is the loss of ability to think for yourself. We're already seeing people are forgetting how to read maps, they're forgetting other skills. If we've lost the ability to be introspective, we've lost human agency and we're spinning around in circles

Impact on Legal System:

- The history of human laws has been built around the assumption that people, not robots, make decisions.
- If robots were shown to have sufficient awareness, then they could be liable as direct perpetrators of criminal offenses, or responsible for crimes of negligence
- A crime consists of two elements: a voluntary criminal act or omission (actus reus) and an intention to commit a crime (mens rea).

- Kingston (2018) references a definition provided by Hallevy (2010) on how AI actions may be viewed under criminal law. According to Hallevy, these legal models can be split into three scenarios

1. Perpetrator-via-another
2. Natural-probable-consequence
3. Direct liability

Identifying who exactly would be held liable for an AI's actions is important, but also potentially difficult.

Impact on Environment and Planet:

1. Use of Natural resources
2. Pollution and waste
3. Energy concerns
4. Ways AI could help the planet

Impact on Trust:

References:

- The Complete Introduction to AI Ethics - Business of Data (business-of-data.com)
- Ethics of Artificial Intelligence and Robotics (Stanford Encyclopedia of Philosophy)
- EPRS_STU(2020)634452_EN.pdf (europa.eu)
- Ethics: Descriptive, Normative, and Analytic (learnreligions.com)
- what is descriptive ethics - Google Search
- Four Branches of Ethics – GKToday
- What are AI Ethics (AI Code of Ethics)? (techtarget.com)
- Ethics vs. Morals – What's The Difference? | Dictionary.com
- https://www.talentica.com/blogs/implementation-of-ai/
- AI Is Transforming Real-Time Data Governance | Simplilearn