Human and artificial intelligence - a critical comparison

Thomas Fuchs

Psychiatric University Clinic Heidelberg

Introduction

„The fact is that AI can go further than humans, it could be billions of times smarter than humans at this point " (Ian Pearson 2018).

„Machines will follow a path that mirrors the evolution of humans. Ultimately, howev-er, self-aware, self-improving machines will evolve beyond humans' ability to control or even understand them " (Ray Kurzweil 2010).

„Singularity" in 2045?

UK
HD

Introduction

Increasingly, human abilities are attributed to artificial systems: perceiving, recognizing, thinking, reasoning, evaluating or deciding.

Conversely, human consciousness appears as a sum of algorithms, a complex data structure in the brain, which in principle could also be realized by electronic systems.

We consider ourselves more and more like our machines, and conversely, we attribute more and more human characteristics to our machines.

UK
HD

Introduction

→ What is the difference between human and artificial intelligence?

Embodied concept of the human person: it is living processes in an organism on which our experience is based.

Without life there is no consciousness.

→ Non-living, technical systems can never attain consciousness, and thus they also lack the decisive prerequisite of human intelligence.
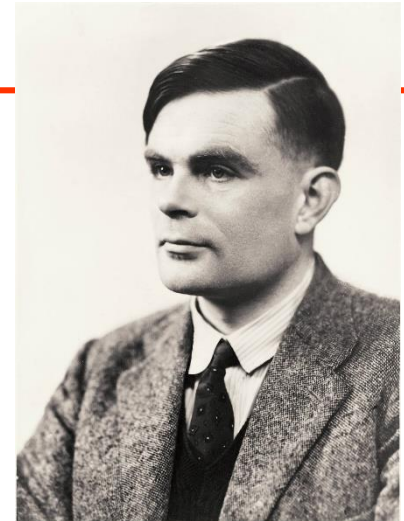
UK
HD

Overview

(1) Subjectivity and simulation

(2) Persons are not programs

(3) Programs are not persons

(4) Problems of the application of AI in judicial decision making

UK
HD

Subjectivity and its Simulation

Turing test (1936): A group of evaluators communicates in writing with a human and with a computer and is then asked to distinguish between human and machine.

Indistinguishability of simulation and original: What acts as intelligently as we do is also intelligent.

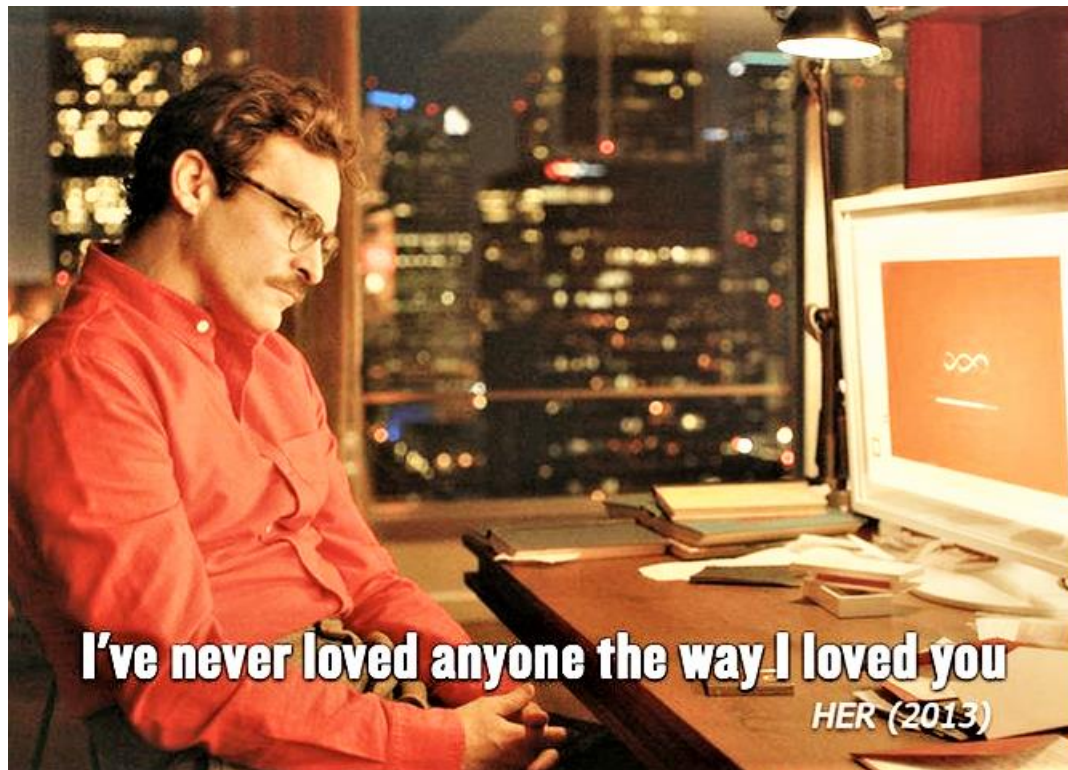„If something looks like a duck, swims like a duck and quacks like a duck, then it is a duck." (??)

UK
HD

Subjectivity and its Simulation

„Sophia", © Hanson Robotics

Subjectivity and its Simulation

„Her" (Spike Jonze, 2013)

I've never loved anyone the way I loved you
HER (2013)

Subjectivity and its Simulation

As Sophia, Alexa, Siri or Samantha show, it is quite possible that we perceive robots and computer systems empathically.

The increasingly perfected simulation is seductive. It demands that we explicitly reject the pretense of subjectivity.

Obviously, we are all too inclined to project our own experience onto the simulations, in a kind of *digital animism.*

UK
HD

Subjectivity and its Simulation

How far does our resistance to simulation extend, how great is its attraction?

When do we give up the distinction between simulation and original?

Is the perfect simulation enough for us in the end?

Subjectivity and its Simulation

— Persons are not programs

— Programs are not persons

(2) Persons are not programs

Functionalism:

Mental states (feelings, perceptions, thoughts, beliefs, etc.) consist of rule-like links between inputs and outputs of a system.

"Pain" is nothing more than the brain state that results in the associated output. This brain state can be described as a certain amount of data.

„Mind is a neural computer, equipped by natural selection with combinatorial algorithms for causal and probabilistic reasoning."

(Steven Pinker 1997)

UK
HD

(2) Persons are not programs

John Searle (1982): The Chinese Room

(2) Persons are not programs

Consciousness is not at all the mindless passing through data states - it is always self-consciousness.

It is for me to feel pain, to perceive, to understand or to think.

Programs and their carrier systems do not experience anything.

The output of such systems is at best the simulation of experience, not the original - what looks, swims and quacks like a duck is far from being a duck.

UK
HD

(2) Persons are not programs

The brain is not a computer:

− no distinction between "hardware" and "software"

− no "data storage": the same thing never happens twice in the brain

− no purely digital signal transmission (neuromodulators)

− embedded in water

UK
HD

(2) Persons are not programs

The brain functions are not conceivable without embodiment:

Consciousness results from vital regulatory processes involving the whole organism, which are integrated in the brainstem and diencephalon.

Bodily self-experience or the feeling of being alive underlies all higher mental functions.

All experiencing is a form of life.

Without life there is no consciousness and also no thinking.

Even a perfect computer simulation of the brain would not be conscious any more than a perfect computer simulation of a hurricane would make us wet or blow us over.

(3) Programs are not Persons

"Artificial intelligence"?

intellegere: "to see, to understand, to grasp".

Intelligence implies seeing oneself and one's situation from a higher perspective = self-consciousness, reflexivity

AI can be described as "…allowing a machine to behave in such a way that it *would be called intelligent* if a human being behaved in such a way" (McCarthy 1956).

$\rightarrow$ The notion of a disembodied intelligence without life and consciousness is self-contradictory. It is only a simulation of narrowly defined areas of human intelligence.
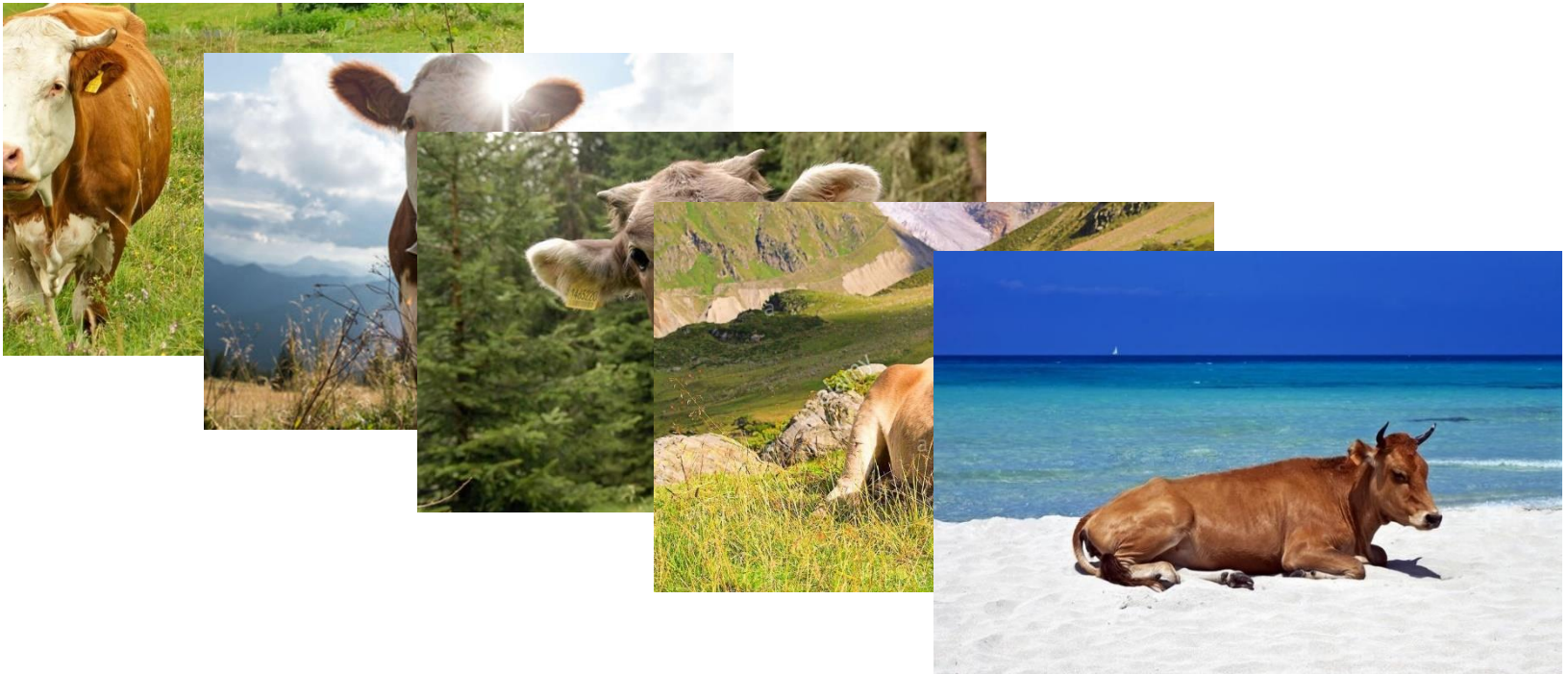
UK
HD

(3) Programs are not Persons

– Do computers solve problems?

– Can computers decide?

– Do "goal-seeking systems" exist?

„Learning systems" (artificial neural networks with progressive adaptation through "training")

UK
HD

(3) Programs are not Persons

Do AI systems recognize objects?

→ No AI systems recognizes anything.

(3) Programs are not Persons

$\rightarrow$ We have been all too hasty in granting the concept of intelligence to our machines.

Summary: Simulation and Original

Intelligence in the true sense is tied to insight, overview and self-awareness: understanding what one is doing.

Prerequisite for consciousness is not only a brain, but a living organism. All experiencing is based on life.

The concept of an unconscious intelligence is a "wooden iron".

However, the fact that we are dealing in AI only with an external-ization of our own calculation and thinking ability, with a *projection* of ourselves, gets increasingly lost to us.

UK
HD

Conclusion: Simulation and Original

Our anthropomorphism tempts us all too easily to ascribe human intentions, actions, even feelings to our machines.

The greatest danger is probably that we voluntarily leave more and more decisions to the systems, which are only transparent to a few and which elude democratic control.

Our supposed artificial doubles are and will remain our products; their intelligence is only the projection of our own.

The decisive challenge of artificial intelligence lies in the question it poses to us and our self-image.

UK
HD

(4) AI in the Judiciary

Auxiliary functions of AI:

- Structured processing and preparation of information

- Preparation of simple judgments

  - Default judgments and statements of inadmissibility

  - E.g. regular divorce, parental authority provision, termination of an employment contract, etc.

UK
HD

(4) AI in the Judiciary

Complex, contradictory cases

— knowledge systems, digital case files

— letting AI prejudge more complex cases?

  — example: increasingly common use of AI systems in the US to assess the recidivism risk of offenders (with an obvious bias against African Americans)

Major problems and risks :

a) Incomprehensibility

b) Datafication

c) Loss of Responsibility

UK
HD

(a) Incomprehensibility

Deep learning techniques lack the explicit logical reasoning or causal inferences that mark conventional human explanations. That result is at odds with equitable justice.

Of course, personal judgment is by no means free of emotional and personal influences. It is also based on juridical and psychological intuitions that may not be completely rationalized.

But the judgment has to justify itself by way of such reasoning, and it remains possible to challenge those opinions through argumentation.

The incomprehensibility of an AI adjudicator could pose legitimacy or fairness problems for individuals who are the subjects of AI adjudication.

UK
HD

(2) Datafication

Deindividualization

AI systems are only based on how other judges have ruled on people with a comparable profile. They are not about the person sitting in front of the judge. However, individuality cannot be defined as a statistical average.

Spreading AI use would necessarily lead to an overly statistical and technological view and a loss of the individual person in the justice system.

Opaque Data Biases

For example, use of crime data that reflects preexisting racial bias in law enforcement

UK
HD

(2) Datafication

Changing view of guilt and culpability

The increasing appeal of AI adjudication could create pressure to adapt the law itself.

For example, moral aspects of the law, such as a requirement that murder occurs only when someone kills with a "malignant heart" would not be accessible to AI information processing.

The legal system could increasingly develop in the direction of a behavioristic social technology for which subjective experience is no longer important.

UK
HD

(3) Loss of Responsibility

Increasing delegation and dissolution of responsibility

As AI plays a larger role in the legal system, human participation will change and decrease.

There are claims that algorithms can already accurately predict court decisions, and we can imagine a legal system without courts as we know them, and where we won't need human judges anymore.

But our entire justice system is decisively based on the fact that legal relations are relations between persons who encounter each other as subjects of freedom, decision and responsibility, and can only accuse, defend, judge and punish each other as such.

UK
HD

(3) Loss of Responsibility

Every jurisdiction is not the execution of a system of rules, which can also be represented by an algorithm, but the establishment and execution of an interpersonal relationship.

The core of this relationship is manifested in the personal encounter between judge and defendant or between judge and litigants.

If this interpersonal relationship is transferred into the execution of a technically driven algorithm, then justice loses its humane basis.

In the end, it would be no more than a social technology.

UK
HD

Conclusion

The more complex society becomes, the more attractive it could become to delegate planning and decision-making to machines.

But no "intelligent system" can tell us what is right, what is good, and what is ethically imperative.

The more the idea of artificial intelligence as a supposedly superior form of analysis, prediction, and evaluation becomes established, the more it might be forgotten that decisions can ultimately only be made by humans themselves.

Responsibility is no technical category.

Conclusion

AI technologies can certainly obscure actual accountability through the supposed objectivity of algorithms.

But in the area of law, it is essentially about human responsibility, both as far as those affected and the judges are concerned.

Justice is based on personal freedom and personal accountability.

If we want to defend the idea of justice, and thus of humanity, we should not hand over our responsibility to algorithms, neither as individuals, nor as societies.

UK
HD

Thank you very much for your attention!

T. Fuchs: In Defense of the Human Being. Foundational Questions of an Embodied Anthropology.
Oxford University Press 2021