[The original, German version of this article appeared on March 20, 2022, as part of the AI column “Aus dem Maschinenraum der KI” in the economics section of the German Sunday newspaper Welt am Sonntag, p. 24.]

[Translated with www.DeepL.com/Translator (free version) - AI technology made in Europe, see https://en.wikipedia.org/wiki/DeepL_Translator, and subsequently polished and modified by the author.]

# No End of Theory in AI!

*Artificial Intelligence thrives on assumptions, theories, and data.*

Do you know Stanley Kubricks' "2001: A Space Odyssey" from 1968? Like "Star Wars" or "Citizen Kane," this film has become a part of our pop culture, and so has the villainous computer HAL. HAL is an artificial intelligence (AI) and is still a prototype of all our hopes and fears about AI: HAL is smart, deceitful, and controls all the systems on the Discovery One spaceship. And it is conscious.

If you haven't heard of HAL, you may be familiar with the famous opening sequence of "2001: A Space Odyssey", in which the sun rises over the earth and moon, accompanied by Richard Strauss' " Thus Spoke Zarathustra" (orig. German title "Also sprach Zarathustra"). This is an allusion to that poetic work of the same title by the German philosopher Friedrich Nietzsche.

AI often makes me think of Nietzsche. Why? "Zarathustra" shows Nietzsche's inclination to extremes. According to the German literary critic Denis Scheck, there is "no mediation or thinking in moderate zones." Valley vs. top of the mountain, strong sv. weak, superhuman (engl. "Übermensch ") vs. the average human. Actually, "Zarathustra" has the subtitle "A Book for All and None".

In the public discourse, AI is also seen either as a savior or as downfall of humanity. This extreme, bipolar view is totally wrong and does address the issue. AI is a science and a tool. Nothing more, but also nothing less. We need a nuanced discussion!

Unfortunately, Nietzsche cannot help us here. He was convinced that there is no causality, and instead life is only an "accidental juxtaposition of things and states". If only Nietzsche were right! I would immediately shave my head! Because, for men, it is easy to observe the high "negative correlation" between the income and the number of hairs on the head: the less hair, the more money.

However, no, really no single visit to the hairdresser has made me richer yet! Men are simply getting older. Statistically, the older, the higher the income. The older, the less hair. Height and weight, on the other hand, are "positively correlated," because they behave in the same way: as you grow, you gain weight. Indeed, this may not be true for everyone of us, but by and large this holds.

Current machine learning algorithms are extremely good at finding correlations. They are less good at telling us about causes and effects: Did this cause that there? What if I did this instead? That is why there is a renewed and strong interest in AI in looking at causal and

logical reasoning. Only this way we can understand the world behind the data and find the fundamental laws of intelligent behavior in humans and machines.

In 2013, Chris Anderson, former editor in chief of Wired magazine, was simply wrong when he heralded "the end of theory" (https://www.wired.com/2008/06/pb-theory/):

> "*This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves.*"

No, you cannot just throw a heap of raw data on a table, press a button, and its value appears. Every data-driven discovery starts already with collecting, processing, cleaning, storing, managing, analyzing, and visualizing the data, and then critically examining the results. Each step in this cycle involves a lot of work and makes assumptions about the data and its value.

There exists no direct route from data to truth! Data are "charged" by assumptions and theories. It was clear to Einstein that his relativity theory had to be verified by qualified experiments. This also holds for AI systems. Ultimately, we must question their predictions through observations in the real world and reject them if necessary. That's how it works in sciences, that's how it works in AI. Sorry Chris!

**Kristian Kersting** is Professor of AI and Machine Learning at TU Darmstadt, co-director of the Hessian Center for AI (hessian.ai), and winner of the "German AI Award 2019". His AI column "Aus dem Maschinenraum der KI" appears regularly in the German Sunday newspaper Welt am Sonntag.