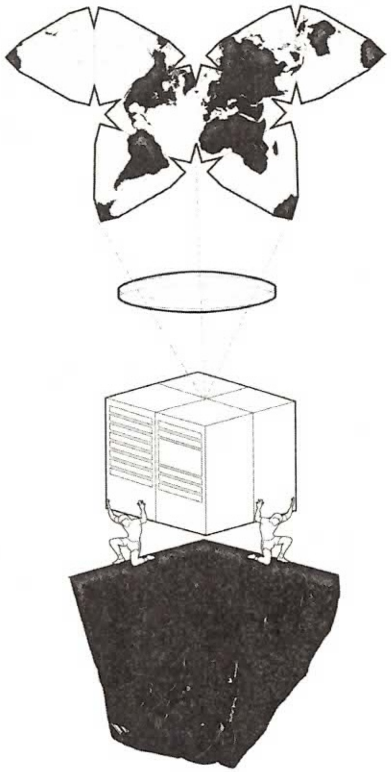
Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.



Introduction

The Smartest Horse in the World

At the end of the nineteenth century, Europe was captivated by a horse called Hans. "Clever Hans" was nothing less than a marvel: he could solve math problems, tell time, identify days on a calendar, differentiate musical tones, and spell out words and sentences. People flocked to watch the German stallion tap out answers to complex problems with his hoof and consistently arrive at the right answer. "What is two plus three?" Hans would diligently tap his hoof on the ground five times. "What day of the week is it?" The horse would then tap his hoof to indicate each letter on a purpose-built letter board and spell out the correct answer. Hans even mastered more complex questions, such as, "I have a number in mind. I subtract nine and have three as a remainder. What is the number?" By 1904, Clever Hans was an international celebrity, with the *New York Times* championing him as "Berlin's Wonderful Horse; He Can Do Almost Everything but Talk."[1]

Hans's trainer, a retired math teacher named Wilhelm von Osten, had long been fascinated by animal intelligence.

Von Osten had tried and failed to teach kittens and bear cubs cardinal numbers, but it wasn't until he started working with his own horse that he had success. He first taught Hans to count by holding the animal's leg, showing him a number, and then tapping on the hoof the correct number of times. Soon Hans responded by accurately tapping out simple sums. Next von Osten introduced a chalkboard with the alphabet spelled out, so Hans could tap a number for each letter on the board. After two years of training, von Osten was astounded by the animal's strong grasp of advanced intellectual concepts. So he took Hans on the road as proof that animals could reason. Hans became the viral sensation of the belle époque.

But many people were skeptical, and the German board of education launched an investigative commission to test Von Osten's scientific claims. The Hans Commission was led by the psychologist and philosopher Carl Stumpf and his assistant Oskar Pfungst, and it included a circus manager, a retired schoolteacher, a zoologist, a veterinarian, and a cavalry officer. Yet after extensive questioning of Hans, both with his trainer present and without, the horse maintained his record of correct answers, and the commission could find no evidence of deception. As Pfungst later wrote, Hans performed in front of "thousands of spectators, horse-fanciers, trick-trainers of first rank, and not one of them during the course of many months' observations are able to discover any kind of regular signal" between the questioner and the horse.[2]

The commission found that the methods Hans had been taught were more like "teaching children in elementary schools" than animal training and were "worthy of scientific examination."[3] But Strumpf and Pfungst still had doubts. One finding in particular troubled them: when the questioner did not know the answer or was standing far away, Hans rarely gave the correct answer. This led Pfungst and Strumpf to con-

Wilhelm von Osten and Clever Hans

sider whether some sort of unintentional signal had been providing Hans with the answers.

As Pfungst would describe in his 1911 book, their intuition was right: the questioner's posture, breathing, and facial expression would subtly change around the moment Hans reached the right answer, prompting Hans to stop there.[4] Pfungst later tested this hypothesis on human subjects and confirmed his result. What fascinated him most about this discovery was that questioners were generally unaware that they were providing pointers to the horse. The solution to the Clever Hans riddle, Pfungst wrote, was the unconscious direction from the horse's questioners.[5] The horse was trained to produce the results his owner wanted to see, but audiences felt that this was not the extraordinary intelligence they had imagined.

The story of Clever Hans is compelling from many angles: the relationship between desire, illusion, and action, the business of spectacles, how we anthropomorphize the nonhuman, how biases emerge, and the politics of intelligence. Hans inspired a term in psychology for a particular type of conceptual trap, the Clever Hans Effect or observer-expectancy effect, to describe the influence of experimenters' unintentional cues on their subjects. The relationship between Hans and von Osten points to the complex mechanisms by which biases find their ways into systems and how people become entangled with the phenomena they study. The story of Hans is now used in machine learning as a cautionary reminder that you can't always be sure of what a model has learned from the data it has been given.[6] Even a system that appears to perform spectacularly in training can make terrible predictions when presented with novel data in the world.

This opens a central question of this book: How is intelligence "made," and what traps can that create? At first glance, the story of Clever Hans is a story of how one man constructed intelligence by training a horse to follow cues and emulate humanlike cognition. But at another level, we see that the practice of making intelligence was considerably broader. The endeavor required validation from multiple institutions, including academia, schools, science, the public, and the military. Then there was the market for von Osten and his remarkable horse—emotional and economic investments that drove the tours, the newspaper stories, and the lectures. Bureaucratic authorities were assembled to measure and test the horse's abilities. A constellation of financial, cultural, and scientific interests had a part to play in the construction of Hans's intelligence and a stake in whether it was truly remarkable.

We can see two distinct mythologies at work. The first myth is that nonhuman systems (be it computers or horses) are analogues for human minds. This perspective assumes that with sufficient training, or enough resources, humanlike intelligence can be created from scratch, without addressing the

fundamental ways in which humans are embodied, relational, and set within wider ecologies. The second myth is that intelligence is something that exists independently, as though it were natural and distinct from social, cultural, historical, and political forces. In fact, the concept of intelligence has done inordinate harm over centuries and has been used to justify relations of domination from slavery to eugenics.[7]

These mythologies are particularly strong in the field of artificial intelligence, where the belief that human intelligence can be formalized and reproduced by machines has been axiomatic since the mid-twentieth century. Just as Hans's intelligence was considered to be like that of a human, fostered carefully like a child in elementary school, so AI systems have repeatedly been described as simple but humanlike forms of intelligence. In 1950, Alan Turing predicted that "at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."[8] The mathematician John von Neumann claimed in 1958 that the human nervous system is "prima facie digital."[9] MIT professor Marvin Minsky once responded to the question of whether machines could think by saying, "Of course machines can think; we can think and we are 'meat machines.'"[10] But not everyone was convinced. Joseph Weizenbaum, early AI inventor and creator of the first chatbot program, known as ELIZA, believed that the idea of humans as mere information processing systems is far too simplistic a notion of intelligence and that it drove the "perverse grand fantasy" that AI scientists could create a machine that learns "as a child does."[11]

This has been one of the core disputes in the history of artificial intelligence. In 1961, MIT hosted a landmark lecture series titled "Management and the Computer of the Future." A stellar lineup of computer scientists participated, including

Grace Hopper, J. C. R. Licklider, Marvin Minsky, Allen Newell, Herbert Simon, and Norbert Wiener, to discuss the rapid advances being made in digital computing. At its conclusion, John McCarthy boldly argued that the differences between human and machine tasks were illusory. There were simply some complicated human tasks that would take more time to be formalized and solved by machines.[12]

But philosophy professor Hubert Dreyfus argued back, concerned that the assembled engineers "do not even consider the possibility that the brain might process information in an entirely different way than a computer."[13] In his later work *What Computers Can't Do,* Dreyfus pointed out that human intelligence and expertise rely heavily on many unconscious and subconscious processes, while computers require all processes and data to be explicit and formalized.[14] As a result, less formal aspects of intelligence must be abstracted, eliminated, or approximated for computers, leaving them unable to process information about situations as humans do.

Much in AI has changed since the 1960s, including a shift from symbolic systems to the more recent wave of hype about machine learning techniques. In many ways, the early fights over what AI can do have been forgotten and the skepticism has melted away. Since the mid-2000s, AI has rapidly expanded as a field in academia and as an industry. Now a small number of powerful technology corporations deploy AI systems at a planetary scale, and their systems are once again hailed as comparable or even superior to human intelligence.

Yet the story of Clever Hans also reminds us how narrowly we consider or recognize intelligence. Hans was taught to mimic tasks within a very constrained range: add, subtract, and spell words. This reflects a limited perspective of what horses or humans can do. Hans was already performing remarkable feats of interspecies communication, public performance, and considerable patience, yet these were not recognized as intelligence. As author and engineer Ellen Ullman puts it, this belief that the mind is like a computer, and vice versa, has "infected decades of thinking in the computer and cognitive sciences," creating a kind of original sin for the field.[15] It is the ideology of Cartesian dualism in artificial intelligence: where AI is narrowly understood as disembodied intelligence, removed from any relation to the material world.

What Is AI? Neither Artificial nor Intelligent

Let's ask the deceptively simple question, What is artificial intelligence? If you ask someone in the street, they might mention Apple's Siri, Amazon's cloud service, Tesla's cars, or Google's search algorithm. If you ask experts in deep learning, they might give you a technical response about how neural nets are organized into dozens of layers that receive labeled data, are assigned weights and thresholds, and can classify data in ways that cannot yet be fully explained.[16] In 1978, when discussing expert systems, Professor Donald Michie described AI as knowledge refining, where "a reliability and competence of codification can be produced which far surpasses the highest level that the unaided human expert has ever, perhaps even could ever, attain."[17] In one of the most popular textbooks on the subject, Stuart Russell and Peter Norvig state that AI is the attempt to understand and build intelligent entities. "Intelligence is concerned mainly with rational action," they claim. "Ideally, an intelligent agent takes the best possible action in a situation."[18]

Each way of defining artificial intelligence is doing work, setting a frame for how it will be understood, measured, valued, and governed. If AI is defined by consumer brands for corporate infrastructure, then marketing and advertising have predetermined the horizon. If AI systems are seen as more reliable or rational than any human expert, able to take the "best possible action," then it suggests that they should be trusted to make high-stakes decisions in health, education, and criminal justice. When specific algorithmic techniques are the sole focus, it suggests that only continual technical progress matters, with no consideration of the computational cost of those approaches and their far-reaching impacts on a planet under strain.

In contrast, in this book I argue that AI is neither *artificial* nor *intelligent*. Rather, artificial intelligence is both embodied and material, made from natural resources, fuel, human labor, infrastructures, logistics, histories, and classifications. AI systems are not autonomous, rational, or able to discern anything without extensive, computationally intensive training with large datasets or predefined rules and rewards. In fact, artificial intelligence as we know it depends entirely on a much wider set of political and social structures. And due to the capital required to build AI at scale and the ways of seeing that it optimizes AI systems are ultimately designed to serve existing dominant interests. In this sense, artificial intelligence is a registry of power.

In this book we'll explore how artificial intelligence is made, in the widest sense, and the economic, political, cultural, and historical forces that shape it. Once we connect AI within these broader structures and social systems, we can escape the notion that artificial intelligence is a purely technical domain. At a fundamental level, AI is technical and social practices, institutions and infrastructures, politics and culture. Computational reason and embodied work are deeply interlinked: AI systems both reflect and produce social relations and understandings of the world.

It's worth noting that the term "artificial intelligence" can create discomfort in the computer science community. The phrase has moved in and out of fashion over the decades and is used more in marketing than by researchers. "Machine learning" is more commonly used in the technical literature. Yet the nomenclature of AI is often embraced during funding application season, when venture capitalists come bearing checkbooks, or when researchers are seeking press attention for a new scientific result. As a result, the term is both used and rejected in ways that keep its meaning in flux. For my purposes, I use AI to talk about the massive industrial formation that includes politics, labor, culture, and capital. When I refer to machine learning, I'm speaking of a range of technical approaches (which are, in fact, social and infrastructural as well, although rarely spoken about as such).

But there are significant reasons *why* the field has been focused so much on the technical—algorithmic breakthroughs, incremental product improvements, and greater convenience. The structures of power at the intersection of technology, capital, and governance are well served by this narrow, abstracted analysis. To understand how AI is fundamentally political, we need to go beyond neural nets and statistical pattern recognition to instead ask *what* is being optimized, and *for whom*, and *who* gets to decide. Then we can trace the implications of those choices.