CHAPTER ONE

CAPITAL, WILL, AND ARTIFICIAL INTELLIGENCE

If one were to start writing a book about artificial intelligence, it might be prudent to ask why we humans consider ourselves an intelligent species. Obviously, there have been many, many answers to this question, but common strands center around such facts as: i) We are technological animals that make and use tools, and we use those tools to continually shape and reshape our environment, or ii) We are political animals that shape and reshape our social structures just as much as we do the rest of our surroundings. This is not a book about artificial intelligence per se (here we write about why people react to AI as they do), but it turns out that some perspective on human intelligence is rather useful nonetheless.

As tool-producing animals, as shapers of environments, we currently have historically unparalleled technological capabilities at our disposal. Any teenager with access to the internet has more information at their fingertips than was contained in the Library of Alexandria, we generate power from splitting atoms, we build bridges of previously unimaginable lengths, and plenty of us have *flown* at least once, sitting in chairs in winged metal tubes. Many of us who do so regularly even forget to marvel at this wonder and instead complain about the lack of leg space, or the quality of the food that we can munch on while flying some 30,000 feet above ground. And of course, we have lately been using advances in statistics to build "learning machines" that

can generate a variety of digital objects, ranging from composite images to plausible responses to text queries—tasks that were, until even more recently, confined to humans.

We also know that in using these very same technologies, we are shaping the environment in ways that are similarly unparalleled. Over the last century, we have been digging and sucking more and more coal and oil out of the planet's crust, burning it to power our technologies. The mere side-effects of using these fossil fuels for energy have been so vast that they may now constitute the basis of a new geological era called the *Anthropocene,* the era of the humans. This era of humanity's own making, is characterized by a rapidly warming climate that poses existential threats to the ecosystems and biological life without which we ourselves could not survive. We continue to produce the greenhouse gasses that drive these changes at an astonishing scale and speed—at the current rate, we will cross the boundary for the relatively "safe" increase of global temperatures of 1.5 °C within about six years. Beyond there be dragons. Awesome powers indeed.

If we have, as toolmakers, gained the ability to destroy ourselves, we might want to ask, why, as political animals, we'd choose to do so. Of course, the question feels ill-posed, somehow. Nobody appears to really be choosing doom. Rather, we have somehow gained the awesome power to change the climate of our whole planet, while simultaneously losing the power *not* to change it. Like in the Disney classic (or Goethe poem) *The Sorcerer's Apprentice*, the technology that we have summoned, has apparently developed a life and will of its own.

In that story, the titular apprentice calls a fleet of simple brooms to life to help him carry buckets of water from a well. He then realizes that he does not know how to stop the brooms, which continue their mindless work, causing a flood. In our story, it is the coal mines and the oil wells, the pipelines and the refineries that we cannot seem to stop. Of course, while the brooms in the story actually spring to life, animated by magic, the mines and wells in the real world have only ever operated through the action of human beings, and none of that has changed. The people who expend their strength and time each day in building and operating these tools, however, are also in some sense not really in control

of them. We all know, after all, that decisions about the amount of oil to be pumped up, the locations at which to drill, etc., are made in boardrooms by a rather different class of people. And like the brooms in the story, such corporate boardrooms operate according to a simple maxim: "*More is always better!*" In the Disney story, more water, in the real world—profit.

In some sense, the will embodied in the machinery of fossil fuel extraction really is simply the will of those at the helm, the CEOs, the board members, the shareholders. And yet, it is an odd kind of will, since no CEO, no board member is properly free to will otherwise. Rather, they are bound by markets, by stock prices, by capitalism. They are bound to seek profit, or else to be cast aside and replaced—a fact that has its psychological correlate in the universal response to ethical conundrums thrown up by the drive for profit: "If we don't do it, someone else will." And indeed, no matter how inadequate this response may seem for a member of this intelligent species of ours, a CEO who produces less than the expected profit will quickly be fired and replaced, and a company that reduces its profit rate quickly loses its market share, and is devalued, possibly up to the point of bankruptcy. At any moment, the capitalist and capital itself can remain what they are only as long as they continue this ceaseless movement towards profit.*

Similarly, capital itself—assets (money, tools, labor) that yield more assets—cannot stop moving, cannot stop growing, or else it will disappear, or be swallowed up by some other entity that can ensure continual movement towards profit. In short, if it stops growing, if it stops chasing profit, it stops *being* capital. When machines are capital, the will embodied in them is the drive of capital towards profit. In this sense, this machine-will is simultaneously somebody's and nobody's. It may appear just as the will of board members, but it is also a structural (and

---

* To put it in terms of Adam Smith—the invisible hand of the market is at the throat even of the capitalists. Picture a Darth Vader—indubitably a master of the invisible hand—as someone who's into autoerotic asphyxiation. A hand more pleasant, surely, than it is for the rest of us, but nonetheless not without dangers.

structuring) pressure of capitalism; one that requires endless movement, demands an endless expansion of assets, profits, capital. To be in power is to embody this *will to profit,* to be merely cast as (en)actor for a role already written in the play of capitalism.

In this way, machines and technology can get imbued with their own motive force, a kind of will and autonomy. This relative autonomy of the technology is, however, a property that technology acquires *as capital*, in its role as an asset that must produce profit. In other words, machines acquire that property not as objects, but as an object in a particular social configuration of humans and things. This relative autonomy of capital isn't only a weird effect of the way our society organizes the economy, it even becomes a socially acceptable—even required—kind of delusion. Corporations are, if not people, at least legal persons, with legal rights—or so the law has decided. And thus, the US Supreme Court famously declared in its 2010 Citizens United ruling, they have the right to free speech. What language do they speak, those legal persons? Well, money is speech, we are told. Don't be fooled into thinking that you can just set this collective delusion aside though—if you find yourself acting as if this didn't make any sense, as if corporations weren't people, you might find yourself subject to rather uncomfortable aspects of the law.

Given this relative autonomy of capital, we must understand technology (both existing and new) not merely on its own terms, but also in its social context, where it is used. The oil well as technology *can* extract oil, the oil well as capital *must* do so. In the same way, we must consider the properties of AI both from the vantage point of the toolmaking animal, and that of the political animal: both as a tool with properties of its own, and as an object that embodies the will of capital.

The cultural critic Mark Fisher wrote about this structural "will" of capitalism that exists so abstractly (casting particular people and machines merely as its bearers) as the *centerlessness of global capitalism*.[1] This centerlessness, he argued, "is radically unthinkable." We cannot help but think that someone somewhere really *is* in charge, really could do otherwise—even if we simultaneously know this to be false. We do not commonly experience wills as abstract and disembodied, casting around for a bearer.

In our ordinary social interactions, a will belongs to a person, and does not ooze out of an economic structure. It is in light of this radically unthinkable centerlessness—the will without a willer—that we have to interpret what happens when a technology emerges that "talks."

A rather interestingly parallel argument about large language models—an AI technology for text generation—was advanced by the linguists and computer scientists Emily Bender, Angelina McMillan-Major, Timnit Gebru, and Margaret Mitchell (at that time, the latter two were working as AI researchers at Google) in a highly influential research paper called *On the Dangers of Stochastic Parrots.*[2] The paper's primary concern is a comprehensive exploration of concrete risks posed by large language models—from environmental impacts to language models encoding negative sentiments against marginalized groups. Google considered the paper dangerous enough to its interests that Gebru and Mitchell were fired shortly after the paper was made public.

In their discussion of interactions between AI and humans, the four of them suggest that humans are predisposed to anthropomorphize large language models due to the way we communicate with each other. When conversing with other humans, we consider our conversation partner's state of mind, intentions, assumptions, beliefs, etc. We employ what psychologists call *theory of mind,* i.e., our capacity to speculate about the contents of other people's minds, and their communicative intent. For example, if we are asked at dinner if we can pass the salt, we interpret this not merely as a question about our ability. We draw inferences about why someone would have asked the question, and conclude that they want salt for their food, and are in fact requesting that we pass it. In other words, engaging with language—in particular, language as a part of dialogue—involves continual, ongoing speculation about the mind behind the words. We cannot help but make these same speculations, no matter the source of the text we encounter. Even if we are using a chatbot, and the text in question actually lacks a human author (having been assembled by an AI algorithm to approximate what text usually looks like), we continue to draw inferences about some communicative intent behind the text.

In other words, just as Fisher's argument is concerned with the unthinkability of the will without willer, Bender et al. are concerned with the unthinkability of text without an intentioned author. In both cases, we are predisposed to act, to think, as if there were an agent behind the things we encounter (events, texts, etc.), regardless of whether we genuinely believe that such an agent actually exists.

It's easy to see how these two tendencies might interact. When we engage in speculations about the future, we tend to quite generally rely on experiences of the past. When we speculate about future autonomous technology, we are prone to extrapolate from our experiences of the only form in which technology is autonomous today—as capital. The will we imagine such future technology to have, will be colored by the "will" we have experienced until today—the will to profit. Now, as technology starts to "talk" to us, speculations about its intentions (communicative and otherwise) will draw on this. The two gaps—the nonexistent willer behind the will, and the nonexistent author behind the text—get identified with each other, as if that made them exist.

We can see this identification at play in an oft-repeated story: the Paperclip AI. For readers who haven't come across this little attempt at a fable/thought experiment, a very brief recap: we are asked to imagine a "superintelligent" AI whose sole aim is the production of paper clips. Guided by a prioritization whose consequences its human makers did not foresee, the AI makes essentially the whole of Earth into a ball of paper clips, dooming humans in their desire for office supplies. In short, the story re-tells the sorcerer's apprentice, and sets it into an explicitly industrial setting. One would think a retelling would clarify the telling we have given above—that it would make clear that the Goethe/Mickey Mouse/Paperclip fable is not simply about tools and the abilities they give us, but about tools as capital, and about the pitfalls of production for the limitless drive to profit. But, alas, the identification of the AI tool with a "superintelligence" in the Paperclip AI version of the fable produces the false identification we sketched above: It is not capital that is understood to be the issue, but "superintelligence." It's not what we do with our tools when directed by the endless will to profit, that appears as the

problem—it's just the tool, whose intelligence, it just so happens, has the very same desire that the boardrooms are cast to enact.

What we will explore in the remainder of the book, then, is based on a suspicion: that speculations about future "super-intelligent AIs" are often, really, speculations about AI models *as capital*. That when we imagine an author behind the artificially generated text, we imagine it as the willer behind the will to profit. That we may be anthropomorphizing, not just a technology, but capitalism itself.

Stories that engage in such anthropomorphization can be powerful, can capture our imagination, precisely because they do reveal something about our very real experiences of life under capitalism. But the angle from which they are told can also obscure things, can make the real mechanisms that threaten us incomprehensible. There is no one willer behind the will-to-profit, there is a structure—that of capitalism, and that of class. After all, the way one experiences the autonomy of capital varies considerably, depending on one's structural position in the system. To be cast in the role of bearer of that will is to be aligned with that will, to be able to embody it. Hence, a capitalist or a CEO may well experience the abstract will as their own. An AI researcher who is wondering if they might get fired for openly discussing technological risks, is likely to experience the drive towards profit as antagonistic to their own autonomy, even if the relation to the technology itself (i.e., their technological expertise) is also the base of their income and relative social power. And of course, many people already experience their jobs as menial and oppressive, and hence the autonomy of capital as diametrically opposed to their own freedom. The stories people tell about future AI can reveal or conceal a lot about the present and the future. Depending on our interpretations, they can clarify something about capitalism, or they can obfuscate, and make reality impossible to see. Which one they do, depends on our ability to understand what about the world these myths and stories makes them resonate—and with who. Or, in other words, it depends on our ability to read stories about artificial intelligences with class on our mind.