WHY WE FEAR AI
ON THE INTERPRETATION OF NIGHTMARES

Hagen Blix and Ingeborg Glimmer

CONTENTS

INTRODUCTION

"The reasons why we have got out of the habit of thinking of our history as progress are obvious. For even when twentieth-century progress is most undeniable, prediction suggests not a continued ascent, but the possibility, perhaps even the imminence, of some catastrophe: another and more lethal world war, an ecological disaster, a technology whose triumphs may make the world uninhabitable by the human species, or whatever current shape the nightmare may take. We have been taught by the experience of our century to live in the expectation of apocalypse."

Eric Hobsbawm, *The Age of Empire* (1989)[1]

A new form of magic is creeping into our world. It is the stirrings of a machine, of an inhuman silicon thing, beginning to act like it is anything but. Through the whirring metal casings that fill the newly overflowing data centers streams not just electricity, but language. For the first time, we can talk to something that is not a human—and that something talks back. Our questions flow into the machines in the data centers, and out flow words, sentences, paragraphs, and pages. At times, it almost feels like there is a someone in that box, not a something, a who and not a what. Whatever it is, a new kind of being is being born. Slowly, day by day, and word by word. Something is animating what was mere transistors before, something called artificial intelligence. A new kind of technological magic? Perhaps.

In science fiction, it is often said that any sufficiently advanced technology is indistinguishable from magic. In other traditions

(literary or otherwise), magic is associated not with technology, but with language, with special words or phrases. Stories are full of spell words, like *abracadabra* or *open sesame,* and when the right person uses them in the right context, they can have powerful effects. In myriad stories, from ancient myths to fantasy novels, simply naming things can bestow power, knowing someone's "true name" means you can control them, and speaking occult languages is associated with special magical talents. Even the ambiguity of the word "spell" betrays magic connotations: one can cast a spell, or one can spell a word (one can even spell trouble or doom).

And indeed, even for those of us who do not believe in magic, language surely has magical qualities. If you make just the right kind of sounds you can initiate a spooky action at a distance, a strange form of mind control. Make some random sounds, and you will attract, at best, strange looks. But if you know just the right "spell," just the right sounds—or just the right way of moving hands and body if your language is signed—and you know just the right recipe to combine them, you can conjure up new thoughts and images in the mind of another. You can transubstantiate the waves of sound or the electromagnetic waves of light into idea-waves. If you have learned to create the right shapes (say, by laying down some pigment on a surface, or chiseling away a stone), you can even talk to people in the future. Admittedly, even if you do master that particular kind of spell(ing), the people from the future cannot talk *back*. But, it certainly sounds like magic otherwise, doesn't it?

Artificial intelligence folds these two ideas of magic into one—it's the first kind, technology, that has begun to use the second kind, language. Up to this point, language was a human specialty, something often taken as a (or even *the*) thing that distinguished our species from all others. Now, it seems, these new machines are wielding their own incantations. What kind of strange thoughts do the machine magicians conjure up in our heads? For some, certainly, it is dreams of all kinds—from personal wealth to somehow solving climate change. But all too often, it seems to be worries, fears, forebodings, and nightmares.

Some of these worries are quite concrete: from misinformation campaigns to the environmental impacts of data centers, from

AI-encoded social biases to the effects of possible automation, AI seems poised to shape how existing social issues develop. In many cases, social conflicts about the path forward are already underway. Labor unions, for example, have been fighting against very concrete AI-driven threats to the livelihoods of their members and to the quality of their work. These worries affect a wide variety of workers, from script writers and actors to longshoremen: in 2023, SAG-AFTRA (the Screen Actors Guild - American Federation of Television and Radio Artists) and the Writers' Guild of America went on strike and won contracts that included restrictions on AI use, mandating that it cannot be used to replace writers,[2] and that using AI simulations of likeness, voice and performance must be consented to and fairly compensated.[3] In 2024, the International Longshoremen's Association also went on strike, demanding limitations on the use of automation, a move that has been widely interpreted as part of a broader pushback against AI.[4]

If one looks around, however, the worries about AI do not stop at threats to ways of making a living. There is, in addition, the stuff of real nightmares, seeping into spaces that are not otherwise known for their flights of fancy. *Time Magazine*, for instance, published a bleak piece on AI development that warned "the most likely result […] is that literally everyone on Earth will die." Illustrating the piece is a gif of a golden circuit diagram: against a flickering black and red background, it sprouts like a tree, circuits reaching out, first wide, then narrow then wide again, until the top blossoms into the shape of a mushroom.[5] Through the flickering colors and the upward and outward growth, it evokes, beautifully and terrifyingly, a specter of a digital nuclear blast. The message is clear.

From humanities professors to billionaires, from AI researchers to spectacular fraudsters, from journalists to the pope, it can seem like all the kinds of people who commonly make it into our media discourses are experiencing equally nightmarish visions. Public statements and open letters warn that humans might be about to lose control or even go extinct. The philosopher of consciousness David Chalmers has said that "in the worst case, [humans are] extinguished entirely, [which] is terrifying."[6] The pop-historian Yuval Harari has teamed up with others to

warn that AI could "hack the foundations of our civilization."[7] Perhaps the briefest of all the open letters is a single sentence: "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."[8] It was signed by, among others, the CEOs of three of the largest AI industry labs, a long list of academics and billionaires, and even former presidents and astronauts. And lest you think we're just hand-picking doom-mongering crackpots on a totally arbitrary basis: at least two of its signatories have since received a Nobel Prize. Crackpots or not, the nightmares certainly are haunting people who are anything but peripheral to today's centers of power.

Eric Schmidt, the former CEO of Google and Alphabet, is also evoking AI nightmares. He has turned into a wandering AI evangelist, a new role involving such stations as Chair of the US National Security Commission on Artificial Intelligence, as well as funding and/or sitting on the board of AI startups. Several of those are defense contractors, selling drones, target identification and/or surveillance systems (coincidentally, precisely the things that the government commission he chaired strongly advocated buying).[9] Schmidt says AI is an "existential risk," which in his words "is defined as many, many, many, many people harmed or killed."[10] Even the Pope, whose thoughts on artificial intelligence are mostly serious reflections on the effects that technological developments will likely have on global inequality, has also called AI "both a terrifying and fascinating tool" that "humans must not lose control of."[11]

AI nightmares may even be inspiring outlandish crime! A former American business executive, Caroline Ellison was involved in what prosecutors called "one of the biggest financial frauds in American history," to the tune of $8 billion. What grand plans motivated her to abscond with that much money? According to her former colleague, "Ellison's drive to make money was the result of her […] desire to use her wealth to stop artificial intelligence from causing humanity's extinction."[12] That sentiment was, apparently, shared by her more famous partner-in-crime, Sam Bankman-Fried, founder of the involved cryptocurrency exchange FTX. He, too, wasn't driven by anything as base as greed,

when they stole billions of dollars from deposits. Bankman-Fried "needed infinity dollars because he planned to address the biggest existential risks to life on earth."[13] Among them: "artificial intelligence that turned on mankind."[14] Their surefire way of rescuing humanity from the pernicious AI menace was… building more AI. They invested at least $500 million in ill-gotten gains in an AI company called Anthropic. Well, what's a couple hundred million stolen dollars if you're having nightmares about getting trapped in the Matrix, or facing the Terminator?

What is this new technology, this alleged "intelligence" that can talk, evoke such nightmares, and elicit such crass money grabs? At its heart, it's a set of statistical techniques, called *machine learning*, that can be used to create mathematical abstractions over data—models. Machine learning algorithms do this by extracting probabilistic patterns from often vast amounts of data. The resulting models can be used to make predictions about new data. Machine learning can be applied to a variety of issues, say, predicting the three-dimensional shapes of proteins from their amino acid sequence, playing Go, identifying objects in images, or producing synthetic speech from text. Useful tools, one might think, and sometimes quite impressive ones—but hardly the stuff of nightmares.

What then gave machine learning methods like "deep learning" a connotation of dark and arcane arts, as if the word "deep" was a reference to raising Lovecraftian "deep ones," and not just a word describing a property of artificial neural networks (perhaps the most prominent tool in the machine learning toolbox, they are called "deep" when their nodes are organized into more than two layers)? Or, to put it another way, why did the mundane statistics of machine learning become almost synonymous with the speculative and even grandiose term "artificial intelligence"? Perhaps one part is simply that applying machine learning techniques to language worked surprisingly well. There are now impressively good models of corpora, of huge amounts of text—language models. Machine learning has produced language models that can match probabilistic properties of the text in corpora, and use those patterns to create contextually plausible continuations of text. In short, machine learning techniques have enabled the

creation of models that can "talk" to us. And yet, this can't be all of it—after all, language in itself is not scary. Someone being able to talk is not generally grounds for special suspicion or extinction nightmares. And computers gaining natural language interfaces in science fiction stories does not universally trigger fear either. Take, for instance, the bridge computer in Star Trek stories—it clearly understands commands of all kinds, from finding and cross-referencing information to setting variables of the environmental controls. And yet, neither viewers nor characters find such interfaces to be particular cause for concern. After all, language may be magic that can put ideas into our heads—but it's not mind control, we aren't forced to believe those ideas.

Perhaps a straightforward, cynical take is the right way to go: maybe both the stoking of fears and branding of statistical tools as "AI" are simply part of that special hyperbole called marketing. After all, if what you're selling could pose an existential risk, it must be very powerful indeed, and maybe its powers could be harnessed for good as well. It calls to mind J.R.R. Tolkien's *The Lord of the Rings*, a story paradoxically beloved by Silicon Valley. Paradoxically, because the ring in that story is indeed powerful— and it must never be used, because its power corrupts. Alas, Silicon Valley and its AI peddlers aren't about to let such "subtleties" dampen the mood—in fact, they love naming their companies for things and characters from Tolkien's Middle Earth. There's Peter Thiel's defense and surveillance company, Palantir (named after a magic stone that allows the user to see other places—though what it shows can be deceptively selective). There's Palmer Luckey's AI, robotics, and defense company, Anduril Industries (named after a sword). Even our wandering AI evangelist Eric Schmidt has funded a Tolkien themed company: Istari (the elvish word for wizards) is another startup that aims to use machine learning for weapons development.[15] A sales pitch that sounds like "look at this powerful thing I'm about to build, it might destroy humanity!" may seem strange—but if you're peddling military equipment, destruction of humans on a large scale is exactly what the customer is looking for.

It's not just AI companies building weapons that deploy such scare-sales tactics though. Sam Altman, the CEO of OpenAI (the

company behind the chatbot ChatGPT) is well-known for trafficking in stark dichotomies. Talking at a 2023 venture capital conference (and hence presumably marketing to possible investors), he described what he saw as the possible future of AI: "I think the best case is like so unbelievably good that it's hard for me to even imagine […] just so unbelievably good that you sound like a really crazy person to start talking about it and the bad case and I think this is like important to say is like lights out for all of us."[16] Why does OpenAI want its investors to believe in unimaginably dangerous AI powers that could either kill all of us or, maybe, be harnessed for good (or at least for a tidy profit)? Well, they are hoping to build and sell, if not the One Ring, then certainly the One Tool. OpenAI is among a handful of companies that intend to sell more than simply some bespoke statistical models, or a better interface for computers that uses text and speech instead of mice and keyboards. No, the ultimate aim—so we're told—is to produce not just artificial intelligence, but artificial *general* intelligence (AGI). What's that? Well, any tool can be said to "surpass" humans in some domain—whether we're talking about a hammer, a microscope, or a calculator, we invented those tools so we could do something that we couldn't do before (or couldn't do as fast, well, painlessly, profitably, etc.). And machine learning-based tools for particular tasks are no exception. But what if there were a tool that's better than people at everything? That's what AGI is supposed to be. In the words of the OpenAI charter—"highly autonomous systems that outperform humans at most economically valuable work."[17] Or in other words—one tool to rule them all.

What better way to make AGI seem imminently achievable than to say the machines are really scary—so scary that they might end up destroying us all (unless the Great Men™ who are building them retain control over them and manage to steer them in the right direction). Certainly, as a marketing and PR move, the scary AGI pitch has a second advantage: if the stakes are as high as the very survival of *humanity*, then concerns about the effects that really existing artificial intelligence technology has right now—from excessive energy usage, to undermining the livelihood of artists and dock workers, to introducing cybersecurity

vulnerabilities, to perpetuating and worsening discrimination, to disinformation and deep fakes, to supercharging mass surveillance, etc., are easily brushed aside.

So, there's certainly good reason for skepticism when we encounter AI fears and nightmares in the media. And yet, even for those of us who think that the hype about AI's future abilities and the scaremongering about its dangers are part of a sales pitch, questions remain. Why do media outlets publish such stories as "The End of Humanity: Will Artificial Intelligence Free Us, Enslave Us—or Exterminate Us" (*The Times*), "AI could go 'Terminator'" (*Fox News*), "AI could pose 'extinction-level' threat to humans and the US must intervene" (*CNN*), or "A.I. May Save Us or May Construct Viruses to Kill Us" (*The New York Times*)?[18] Perhaps more importantly: what makes such stories resonate—especially with those of us who own neither crypto-exchanges to defraud people with, nor chatbots or automated weaponry that they want to sell? What do the nightmares of the world's Eric Schmidts and crypto-fraudsters have to do with the worries of the union workers?

Maybe the OpenAI definition of AGI can provide us with a crucial hint—particularly the "economically valuable work" part. The science fiction author Ted Chiang remarked in an interview that, perhaps, "most fears about A.I. are best understood as fears about capitalism […] as fears or anxiety about how capitalism will use technology against us."[19] If that is true, then the nightmares of billionaires like Eric Schmidt and those of the union workers must be about rather different things. After all, Eric Schmidt is far more likely to find himself on the side of those who do the using, rather than on the side of those that the technology might be used against. That is, of course, the nature of capitalism: some people own vast amounts of assets (e.g., shares in companies, real estate, patents, and, of course, technology), and they use those assets to make more money, and more assets. The rest of us, those who do not own assets that can make more assets (or *capital*), make a living by working for those people and institutions that do. Whether it's a forklift, or a computer cluster—capitalists use their ownership and control over assets, over technology, to direct the work that the rest of us do for a wage, so that they can make a profit. In

other words, capitalism is a system with social classes, and what class someone belongs to changes how they relate to technology.

If we want to understand AI nightmares, then, we must interpret them with class in mind. Perhaps even the very same myths about AI doom may turn out to ring true to different people for quite different reasons—we'll each have to interpret our own nightmares. We'll have to figure out who gets to be considered part of that vague and elusive "we," of "humanity," that future AI supposedly threatens, and who is merely an afterthought. We'll have to make sure not to get taken in by the fables of the crypto-fraudsters or the weapons peddlers (some of whom already weave the very nightmares that we live in). Let's get into the thick of the politics and economics of AI—and interpret AI nightmares against that backdrop. Along the way, we'll encounter an array of colorful characters from autonomous technology with its own will, to steam demons, from technological deities to Friedrich Hayek's ghost. We'll see how capitalism and class shape both the actual tools called AI and the AI nightmares. And of course, we'll try to find, among the nightmares, inspiration for a better future.

PART I

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.