**Silicon Valley is Turning into Its Own Worst Fear**

**BuzzFeedNews, Ted Chiang,**

**The importance of metacognition: taking a step back and reflecting on and re-evauating what we are doing: the consequences and unforeseen consequences.**

**In psychology,** **the term “insight”** is used to describe a recognition of one’s own condition, such as when a person with mental illness is aware of their illness. More broadly, it describes the ability to recognize patterns in one’s own behavior. It’s an example of metacognition, or thinking about one’s own thinking, and it’s something most humans are capable of but animals are not. And I believe the best test of whether an AI is really engaging in human-level cognition would be for it to demonstrate insight of this kind.

Insight is precisely what Musk’s strawberry-picking AI lacks, as do all the other AIs that destroy humanity in similar doomsday scenarios. I used to find it odd that these hypothetical AIs were supposed to be smart enough to solve problems that no human could, yet they were incapable of doing something most every adult has done: taking a step back and asking whether their current course of action is really a good idea. Then I realized that we are already surrounded by machines that demonstrate a complete lack of insight, we just call them corporations. Corporations don’t operate autonomously, of course, and the humans in charge of them are presumably capable of insight, but capitalism doesn’t reward them for using it. On the contrary, capitalism actively erodes this capacity in people by demanding that they replace their own judgment of what “good” means with “whatever the market decides.”

Because corporations lack insight, we expect the government to provide oversight in the form of regulation, but the internet is almost entirely unregulated. Back in 1996, John Perry Barlow published a manifesto saying that the government had no jurisdiction over cyberspace, and in the intervening two decades that notion has served as an axiom to people working in technology. Which leads to another similarity between these civilization-destroying AIs and Silicon Valley tech companies: the lack of external controls. If you suggest to an AI prognosticator that humans would never grant an AI so much autonomy, the response will be that you fundamentally misunderstand the situation, that the idea of an ‘off’ button doesn’t even apply. It’s assumed that the AI’s approach will be “the question isn’t who is going to let me, it’s who is going to stop me,” i.e., the mantra of Ayn Randian libertarianism that is so popular in Silicon Valley.

The ethos of startup culture could serve as a blueprint for civilization-destroying AIs. “Move fast and break things” was once Facebook’s motto; they later changed it to “Move fast with stable infrastructure,” but they were talking about preserving what they had built, not what anyone else had. This attitude of treating the rest of the world as eggs to be broken for one’s own omelet could be the prime directive for an AI bringing about the apocalypse. When Uber wanted more drivers with new cars, its solution was to persuade people [with bad credit](https://www.buzzfeed.com/tedchiang/the-real-danger-to-civilization-isnt-ai-its-runaway?utm_term=.fndXk5EL3) to take out car [loans](https://www.buzzfeed.com/tedchiang/the-real-danger-to-civilization-isnt-ai-its-runaway?utm_term=.fndXk5EL3) and then deduct payments directly from their earnings. They positioned this as disrupting the [auto loan](https://www.buzzfeed.com/tedchiang/the-real-danger-to-civilization-isnt-ai-its-runaway?utm_term=.fndXk5EL3) industry, but everyone else recognized it as predatory lending. The whole idea that disruption is something positive instead of negative is a conceit of tech entrepreneurs. If a superintelligent AI were making a funding pitch to an angel investor, converting the surface of the Earth into strawberry fields would be nothing more than a long overdue disruption of global land use policy.

There are industry observers talking about the need for AIs to have a sense of ethics, and some have proposed that we ensure that any superintelligent AIs we create be “friendly,” meaning that their goals are aligned with human goals. I find these suggestions ironic given that we as a society have failed to teach corporations a sense of ethics, that we did nothing to ensure that Facebook’s and Amazon’s goals were aligned with the public good. But I shouldn’t be surprised; the question of how to create friendly AI is simply more fun to think about than the problem of industry regulation, just as imagining what you’d do during the zombie apocalypse is more fun than thinking about how to mitigate global warming.

There have been some impressive advances in AI recently, like AlphaGo Zero, which became the world’s best Go player in a matter of days purely by playing against itself. But this doesn’t make me worry about the possibility of a superintelligent AI “waking up.” (For one thing, the techniques underlying AlphaGo Zero aren’t useful for tasks in the physical world; we are still a long way from a robot that can walk into your kitchen and cook you some scrambled eggs.) What I’m far more concerned about is the concentration of power in Google, Facebook, and Amazon. They’ve achieved a level of market dominance that is profoundly anticompetitive, but because they operate in a way that doesn’t raise prices for consumers, they don’t meet the traditional criteria for monopolies and so they avoid antitrust scrutiny from the government. We don’t need to worry about Google’s DeepMind research division, we need to worry about the fact that it’s almost impossible to run a business online without using Google’s services.

It’d be tempting to say that fearmongering about superintelligent AI is a deliberate ploy by tech behemoths like Google and Facebook to distract us from what they themselves are doing, which is selling their users’ data to advertisers. If you doubt that’s their goal, ask yourself, why doesn’t Facebook offer a paid version that’s ad free and collects no private information? Most of the apps on your smartphone are available in premium versions that remove the ads; if those developers can manage it, why can’t Facebook? Because Facebook doesn’t want to. Its goal as a company is not to connect you to your friends, it’s to show you ads while making you believe that it’s doing you a favor because the ads are targeted.

So it would make sense if Mark Zuckerberg were issuing the loudest warnings about AI, because pointing to a monster on the horizon would be an effective red herring. But he’s not; he’s actually pretty complacent about AI. The fears of superintelligent AI are probably genuine on the part of the doomsayers. That doesn’t mean they reflect a real threat; what they reflect is the inability of technologists to conceive of moderation as a virtue. Billionaires like Bill Gates and Elon Musk assume that a superintelligent AI will stop at nothing to achieve its goals because that’s the attitude they adopted. (Of course, they saw nothing wrong with this strategy when they were the ones engaging in it; it’s only the possibility that someone else might be better at it than they were that gives them cause for concern.)

**There’s a saying,** popularized by Fredric Jameson, that it’s easier to imagine the end of the world than to imagine the end of capitalism. It’s no surprise that Silicon Valley capitalists don’t want to think about capitalism ending. What’s unexpected is that the way they envision the world ending is through a form of unchecked capitalism, disguised as a superintelligent AI. They have unconsciously created a devil in their own image, a boogeyman whose excesses are precisely their own.

Which brings us back to the importance of insight. Sometimes insight arises spontaneously, but many times it doesn’t. People often get carried away in pursuit of some goal, and they may not realize it until it’s pointed out to them, either by their friends and family or by their therapists. Listening to wake-up calls of this sort is considered a sign of mental health.

We need for the machines to wake up, not in the sense of computers becoming self-aware, but in the sense of corporations recognizing the consequences of their behavior. Just as a superintelligent AI ought to realize that covering the planet in strawberry fields isn’t actually in its or anyone else’s best interests, companies in Silicon Valley need to realize that increasing market share isn’t a good reason to ignore all other considerations. Individuals often reevaluate their priorities after experiencing a personal wake-up call. What we need is for companies to do the same — not to abandon capitalism completely, just to rethink the way they practice it. We need them to behave better than the AIs they fear and demonstrate a capacity for insight. ●