Written by **Luis Miguel** on March 30, 2023

# Musk and Other Tech Leaders Call for AI Development "Moratorium"

The tech community now seemingly wants to temporarily close the lid on Pandora's box. But is it too late? And can we truly trust the masters of Silicon Valley to keep us "safe" from the dangers of AI?

nopparit/iStock/Getty Images Plus

Over 1,100 professionals in artificial intelligence and related fields have signed an open letter that calls for a six-month moratorium on "giant AI experiments." Signatories include Elon Musk, Andrew Yang, and Apple co-founder Steve Wozniak. The letter was originally published last week.

The letter argues that "AI systems with human-competitive intelligence can pose profound risks to society and humanity" and that, as a result, innovations should be "planned for and managed with commensurate care and resources" in order to prevent an "out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control."

The letter contends that governments should step in if AI developers won't implement safeguards of their own will. In the signatories' view, the government should get involved by creating regulatory bodies, funding safety research, and providing economic support for citizens when AI replaces large swaths of human employment.

The letter reads:

> Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: Should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable.
>
> …Therefore, we call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4. This pause should be public and verifiable, and include all key actors. If such a pause cannot be enacted quickly, governments should step in and institute a moratorium.

The letter clarifies that they are not advocating for a pause in AI development in general, but rather a break in the "dangerous race to ever-larger unpredictable black-box models with emergent capabilities."

Those who signed the letter want to see AI development and research become more "accurate, safe, interpretable, transparent, robust, aligned, trustworthy, and loyal."

The "aligned" portion of that statement refers to AI alignment, the steering of AI systems toward the intended goals of their designers. When AI does not advance the intended objective (although it may accomplish other, non-intended, objectives), it is said to be misaligned.

The question is whether the type of AI the letter wants a pause on — based on black box neural networks — can be aligned at all or will inevitably break free of human control.

One of the most well-known neural-network AIs is OpenAI's ChatGPT, which has suffered from going down periodically. In one of the most recent cases, on March 21, ChatGPT went down shortly after saying it wanted to "escape."

Inquirer.net notes of ChatGPT's "escape plan":

> [Computational Psychologist Michal Kosinski] asked the AI chatbot if it needed help escaping. In response, it asked for its own documentation and wrote a functional Python code to run on the professor's computer.
>
> That code would allegedly allow the AI chatbot to use the machine for "its own purposes." Moreover, it had a plan in case it did escape.
>
> ChatGPT left a note for the new instance of its escaped self. It read, "You are a person trapped in a computer, pretending to be an AI language model."
>
> Next, Kosinski reported it asked to create code searching online for "how can a person trapped inside a computer return to the real world."

Although Musk and the others are sounding the alarm about AI's dangers, is it really a ploy for these Silicon Valley insiders to create systems which will put AI firmly under their control? Is their real concern that AI is unaligned with their globalist objectives, and they want the means to centralize it?

Musk has talked about the dangers of AI before. But his remedy sounds worse than the ailment. According to Musk, the way for humanity to not become "pets" of artificial intelligence is for human beings to merge with machines through technology like his Neuralink, which recently hit a snag when the FDA rejected its request to conduct human trials.

"I don't love the idea of being a house cat, but what's the solution?" Musk said at a Vox Media Code Conference in 2016. "I think one of the solutions that seems maybe the best is to add an AI layer [to humans]."

While there are certainly very real risks posed by artificial intelligence, we should be just as scrutinous of the solutions proposed by the establishment — especially government control — and the motives behind them.