

# AI Answers Now

September 2023 • [press@futureoflife.org](mailto:press@futureoflife.org)

AI corporations are locked in an out-of-control arms race to develop more and more powerful AI systems, with no clear plan to make them safe or keep them under control. These developers acknowledge that their systems endanger human lives today - from AI-enabled bio-weapons to massive cyber attacks - and could make us extinct. Yet they are not slowing down, and have been unwilling or unable to guarantee our safety. In doing so they are jeopardizing our shared future, including the incredible benefits AI could bring if developed carefully and responsibly.

These companies must demonstrate that they understand the products they are creating and the threats they represent, and that they are taking steps to guarantee our safety. They must answer the questions below:

## Critical Questions

1. **Could the systems you are building destroy civilization?** Is it a possibility that AI could grow powerful enough to destroy or permanently disempower humanity?
2. **Do you understand how these systems work?** Do you fully understand how current systems work, and how they make decisions? If you do not, doesn't this present a problem to ensuring they are safe?
3. **What would make you pause, and how would you do it?** Many AI researchers have said they'll pause developing more powerful AI if the risk grows too great. Under what circumstances would you pause, and how can we ensure that this threshold doesn't simply get moved if it is approached? How would you actually implement such a slowdown?
4. **Could your systems be used to kill and hurt people?** AI labs have admitted that advanced AI could significantly enable bio-terror attacks in the next 2-3 years. How will you prevent misuse by bad actors?
5. **Are you responsible for harms?** Current AI systems have already harmed society. Are you responsible for the harms caused by the products you develop? Are you liable?
6. **Could we lose control?** Do you take seriously the possibility that humanity may lose control over advanced AI systems? What is your plan for preventing this?
7. **Is it democratic?** Given the level of power you ascribe to these systems, how is an advanced AI compatible with democracy? Will you destabilize democracies by giving individual agents such terrifying power?

# Statements from AI Corporations

## Statement from the Center for AI Safety

- “Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.” (statement [signed](#) by OpenAI CEO Altman, DeepMind CEO Hassabis, Anthropic CEO Amodei, Turing Prize winners Hinton and Bengio, May 2023)

## OpenAI CEO Sam Altman

- “Development of superhuman machine intelligence (SMI) is probably the greatest threat to the continued existence of humanity.” ([Sam Altman’s blog](#), Feb 2015)
- “A misaligned superintelligent AGI could cause grievous harm to the world; an autocratic regime with a decisive superintelligence lead could do that too.” ([Open AI’s website](#), Feb 2023)

## DeepMind CEO Demis Hassabis

- “When it comes to very powerful technologies—and obviously AI is going to be one of the most powerful ever—we need to be careful,” he says. “Not everybody is thinking about those things. It’s like experimentalists, many of whom don’t realize they’re holding dangerous material.” ([Time](#), Jan 2023)

## DeepMind Co-founder and Chief Scientist Shane Legg

- “We’re not going to have a practical theory of friendly AI. I’ve spoken to a bunch of people [...] none of them, that I’ve ever spoken to, think they will have a practical theory of friendly AI in about ten years time. No way. [...] It’s really really hard. We have no idea how to solve this problem.” ([Machine Super Intelligence](#), Aug 2009)

## Google CEO Sundar Pichai

- “We don’t have all the answers there yet - and the technology is moving fast [...] So does that keep me up at night? Absolutely.” ([Sky News](#), April 2023)

## Anthropic CEO Dario Amodei

- “We found that today’s AI systems can fill in some of these steps, but incompletely and unreliably – they are showing the first, nascent signs of risk. However, a straightforward extrapolation of today’s systems to those we expect to see in 2-3 years suggests a substantial risk that AI systems will be able to fill in all the missing pieces, if appropriate guardrails and mitigations are not put in place. This could greatly widen the range of actors with the technical capability to conduct a large-scale biological attack.” ([Testimony to the US Senate](#), July 2023)

## Useful Data

- Half of A.I. experts think human inability to control AI has a 10% chance of causing human extinction or similarly permanent and severe disempowerment of humanity ([AI Impacts survey](#)).
- 82% of Americans don't trust tech executives to regulate AI. 63% of Americans want regulation to actively prevent superintelligent AI ([AIPI/YouGov](#)).
- 67% of voters agreed we should restrict how powerful AI models can become even though this risks making American companies fall behind China, and only 14% disagreed ([AIPI/YouGov](#)).
- 86% believe AI could accidentally cause a catastrophic event.
- 72% prefer slowing down the development of AI compared to just 8% who would rather speed it up.
- Three in four Democrats and Republicans alike believe artificial intelligence could eventually pose a threat to the existence of the human race.
- 54% believe that human-level AI (AGI) will be developed within 5 years.
- 36% of AI researchers believe that AI could cause a "nuclear-level catastrophe" ([Stanford Report](#)).
- 55% of Americans think AI could eventually pose an existential threat. ([Monmouth University poll](#)).
- A global study of AI attitudes found that, across 17 major countries, 71% believe AI regulation is necessary ([KPMG](#)).
- 55% of Americans are in favor of "having a federal agency regulate the use of artificial intelligence similar to how the FDA regulates the approval of drugs and medical devices" ([Monmouth University poll](#)).
- 69% of Americans support a six-month pause in AI development. 4% strongly oppose ([YouGov poll](#)).