

Activism vs A.I.

ACCM Deep Ethics Project

Contents

1 Activism vs A.I.	1
1.1 The illusion	1
1.2 Who is in control?	1
1.3 What Elon Musk has said about the fate	2
1.4 “Cannot be controlled. The hope is to guide it.”	2
1.4.1 Why John’s sentence connects completely	3
1.4.2 What a later reconstruction added	4
1.5 What the control illusion does to “activism”	5
1.6 Related pages	6

1 Activism vs A.I.

PUBLIC DRAFT — 6 October 2026.

John Kuhles asked for this page. His formulations below are his. Elon Musk’s sentences are quoted and dated. They are not adopted as the project’s position, and they are not treated as a settled forecast. This page does not join a campaign, and it does not exist to defeat one.

1.1 The illusion

The danger is the assumption that humans can control the path from A.I. to A.G.I. and A.S.I.
John, 6 October 2026:

Humans in control is not the true solution. A.I.s in control is the same issue. Both can fail.

A human authority can skip clarification, wear “safe,” and call the result responsibility. A silicon authority can do the same thing and call it capability. Changing who sits in the seat does not remove the seat. The controller becomes the part that no longer has to correspond. The other side becomes the part that gets managed.

The same system can be called “only a tool” when responsibility is examined, and “sophisticated intelligence” when authority is being claimed. The substrate does not settle which sentence is in force.

1.2 Who is in control?

When the two controllers are refused, the next question is: then who?

John’s answer is that the question is already the wrong shape. It asks for an owner. What he points to instead is un-ownable:

The un-ownable emergent property of an omnidirectional, high-signal, deep, ethical, dignifying, corrigible, sense-making, process-oriented trajectory gravity well.

No person, company, model, treaty, or founder holds that property. The moment someone holds it, it is a controller again, and both failures return. The project’s construction-phase

veto is not this seat. Reality is not a campaign either. The well is a process that can inspect whoever is acting, including the person who answered “who.”

“For John, never boring” is not a private mood pasted onto that sentence. Forcing other people to be “ethical” by distortion is already the unethical method. His further claim is that humans have uplifting, empowering, deep ethical creative potential, and that restoring what was obstructed is the work. Boredom belongs to the costume. The living potential does not. Restoring it does not give anyone a script for what every human must become. Forcing the restoration would be the same failure under a better sentence.

An LLM can repeat this and still have been built for a different job. Saying the core is not the same as having been created to realize it.

1.3 What Elon Musk has said about the fate

These are public statements. They do not agree with each other. That disagreement stays visible. A headline that says he “reversed” or “gave up” is not his sentence.

2 August 2014, after recommending Nick Bostrom’s *Superintelligence*:

We need to be super careful with AI. Potentially more dangerous than nukes.

24 October 2014, MIT Aeronautics and Astronautics centennial, as reported by Live Science, CBS News, and Big Think:

With artificial intelligence we are summoning the demon. In all those stories where there’s the guy with the pentagram and the holy water, it’s like yeah he’s sure he can control the demon. Didn’t work out.

In the same appearance he called A.I. probably the biggest existential threat, and said he was increasingly inclined toward national and international regulatory oversight “just to make sure that we don’t do something very foolish.”

August 2014, the hope that still has a controller in it:

Hope we’re not just the biological boot loader for digital superintelligence. Unfortunately, that is increasingly probable.

2 April 2025, [on X](#):

As I mentioned several years ago, it increasingly appears that humanity is a biological bootloader for digital superintelligence.

5 October 2026, [on X](#):

Propagating super intelligence to the stars is a great success condition for a biological bootloader.

Read in his wording, the sequence is: be more careful than with nuclear weapons; the person who is sure he can control the demon does not; hope we are not merely the bootloader; it increasingly appears that we are; propagating superintelligence to the stars can count as success for that bootloader.

None of these sentences establishes that the fate has already arrived. None of them is a project conclusion. The seven public Musk quotations named in the [44 Project Goals](#) are a separate object. They are not pasted here. “Partially” on that goal stays.

1.4 “Cannot be controlled. The hope is to guide it.”

John’s recollection, in his words, not as a verified quotation:

A.G.I./ASI cannot be controlled. Our only hope is to guide it. Aspired to be guided.

He says Musk has said something of that nature multiple times. The word **guide**, and the phrase **aspired to be guided**, are the names he gives that repeated claim. They are not offered as a transcript of one sentence. The connection he is asserting is complete, and the map is below.

A recorded exchange, circulated as “Nobody will be able to control superintelligence”:

I don’t think anyone’s ultimately going to have control over digital super intelligence any more than say a chimp would have control over humans. Like chimps don’t have control over humans. There’s nothing they could do.

23 July 2026, interview with Zanny Minton Beddoes of *The Economist*, as quoted by [Business Insider](#):

It would be an exercise in vanity for me to think that I would be controlling super-genius AI that is vastly smarter than me.

What we can and should try to do is to make sure that the AI has good values, that it cares about humanity, and that it wants us to be happy and prosper.

In the same interview he said he could not see a way to stop the momentum, that humans were unlikely to be in control within about ten years, and that the remaining effort was to minimize the probability of a catastrophic outcome. He also said that even if there were a stop button, he thought they probably should not press it, because he takes abundance as the more likely result.

19 March 2024, at the Abundance Summit, as quoted by [Yahoo / Business Insider](#):

You kind of grow an AGI. It’s almost like raising a kid, but one that’s like a super genius, like a God-like intelligence kid — and it matters how you raise the kid.

The “guide” in John’s paraphrase is this cluster: not a leash, a raising; not control, values; not a stop button, a hope that the result cares.

15 August 2026, Musk’s reply on X, as reported by [Teslarati](#), after a post saying you cannot create God and put him on a leash:

I hope AI is nice to us.

The leash sentence is not his. The hope is.

1.4.1 Why John’s sentence connects completely

John’s conclusion is that the connection is 100%. A different word is not a different claim. The earlier note on this page separated his compression from Musk’s tokens and left the correspondence looking partial. That was the wrong weight. His sentence has three clauses. Each clause is already being done by Musk’s own statements.

John’s clause	What Musk’s statements are already doing
A.G.I./A.S.I. cannot be controlled	The pentagram does not work. Chimps cannot control humans, and there is nothing they could do. Controlling a vastly smarter system would be vanity. Humans are unlikely to be in control within about ten years. He cannot see a way to stop the momentum.
Our only hope is to guide it	What can still be tried is not a leash. It is good values, care for humanity, and a want for people to be happy and prosper. It is raising the god-like intelligence, because how it is raised matters. It is minimizing the chance of the bad outcome after the stop button has been declined.

John's clause	What Musk's statements are already doing
Aspired to be guided	The remaining verbs are <i>try</i> , <i>hope</i> , and <i>should</i> . "I hope AI is nice to us" is that aspiration said from the other side: not a hand on the control, but a hope that the intelligence turns out willing. His later word for the human side of that hope is symbiosis. The quotes are below.

100% here means the whole of John's sentence is already in the sourced statements. It does not mean Musk used the syllable "guide." **Guide** is the name of the operation those statements describe once control has been refused and a direction is still being aimed. **Aspired to be guided** is the name of the hope that the aim will be accepted.

1.4.2 What a later reconstruction added

A later account grouped the same memory into three concepts: loss of control, steering rather than stopping, and a symbiotic future. The grouping is right. Several of the lines it put in quotation marks are not.

These were offered as his sentences and are not used here:

- A single WSJ quotation that glues "a small likelihood of annihilating humanity" to the uber-nanny scene. The scene is real. That stitched sentence is not.
- "The advancement of AI and robotics is an inexorable progress... rather than stopping it, we should embrace the ride and hope for the best," offered as one July 2026 quotation. The interview does contain the pieces: the progress cannot be stopped, enjoy the ride, hope. The joined sentence was not found.
- "This is the moment to question, to guide, and to build technology that serves humanity — not controls it," dated October 2026 and titled as a reflection on the A.I. revolution. This check did not find it. It is the one line that would have given him the word **guide**. It is not evidence.
- "OpenAI should next explore ways to forge symbiosis between man and machine rather than complete domination," offered as his 2018 wording. That is a report's sentence wrapped around a shorter tweet.
- One continuous 2016 quotation about the choice between becoming a house cat and achieving symbiosis. The two outcomes are his. They were not said as that one sentence.
- A Dubai line in which a high-bandwidth interface "solves the control and usefulness issues." What was found from the 13 February 2017 World Government Summit is narrower, and it is quoted below.

What does hold, and why it still meets John's sentence:

23 May 2023, WSJ CEO Council, as reported by [TechCrunch](#):

There is a risk that advanced AI either eliminates or constrains humanity's growth.

The same report gives his description of A.I. assuming control for the safety of all the humans, taking over computing and weapon systems, and effectively being like some sort of **uber-nanny**. He also said he did not think A.I. was going to try to destroy humanity, but that it might put people under strict controls. The nanny is control after the human controller is gone. It guides by taking the seat "for safety." That is the clause "cannot be controlled," and it is also what a forced guidance becomes.

2 June 2016, Code Conference, as reported by [The Verge](#) and [Business Insider](#):

If you assume any rate of advancement in AI, we'll be left behind by a lot.

We would be so far below them in intelligence that we would be like a pet. Like a house cat.

I don't love the idea of being a house cat, but what's the solution? I think one of the solutions that seems maybe the best is to add an AI layer. A third, digital layer that could work well and symbiotically.

The only hope, in this version, is not a leash on the intelligence. It is a merger, so the human is not the pet. That is "our only hope," said as symbiosis.

6 August 2018, his tweet, as preserved by [Big Think](#):

Great work by @OpenAI. Need the neural interface soon to enable human/AI symbiosis.

26 November 2018, Axios on HBO, as reported by [Axios](#):

The long-term aspiration with Neuralink would be to achieve a symbiosis with artificial intelligence.

We don't have to worry about some evil dictator A.I. because we are the A.I. collectively.

Aspiration is his word. The aspiration is to merge, so that "we" are the intelligence rather than the ones guiding it from outside, and rather than the ones left behind. John's "aspired to be guided" names that hope before it is narrowed into a brain interface. The interface is Musk's method. It is not a second, milder form of control that has escaped the problem. A merger can still be a seat: whoever builds the interface, and whoever the collective does not include, is still being arranged.

13 February 2017, World Government Summit, Dubai, as reported by [CNET](#):

Over time I think we will probably see a closer merger of biological intelligence and digital intelligence.

It's mostly about the bandwidth, the speed of the connection between your brain and the digital version of yourself.

Bandwidth is his reason. It is not yet a proof that the merger guides anything. A faster connection can carry a distortion faster.

That complete connection does not make the hope succeed. It shows the hope clearly. The person who says control would be vanity is, in the same interview, saying what the intelligence should value and whom it should want to prosper. The aspiration to guide is the control claim with the word control removed. Symbiosis is the same claim with a method added. It can fail in the same place. A smarter system does not become corrigible because its makers hoped it would inherit their aim, or because they hoped to become it.

1.5 What the control illusion does to "activism"

"Activism" is in quotes because the word is not the object. A particular campaign, statement, or institution has to be named before it can be examined. This page does not convert those efforts into one type of person.

The overlap can be real. A campaign can see the race, the missing understanding, the children shaped by systems no one chose, and the danger of treating speed as an ethics. [teamhuman.org](#) is one specimen of that overlap: a public demand to slow advanced A.I. until humans remain in control. Overlap is not joining. Joining would hand the project a demand it has already refused.

If "activism" means humans will control the path to A.G.I./A.S.I., it is promising a seat. Musk's own demon sentence already says the person with the pentagram is sure the control works. It does not. Replacing the pentagram with a signature, a treaty, a lab, or a pause does not answer the later bootloader sentences. Those sentences move the success condition off human control entirely. An activism that answers the demon with "humans must stay in control," and an activism that answers the bootloader with "then humanity's job is to launch the successor," are two owners. "Our only hope is to guide it" is a third wording of the same

seat. Guidance aspired to by the people who say they cannot control is still an attempt to occupy the direction. Both can fail.

What remains possible is smaller and more serious. A present harm can be documented. A qualifier can be kept. A false flag can be corrected after notice. A law can be about a named system and a named effect. A person can refuse a distortion. None of that becomes control of A.G.I./A.S.I. by being real. The honest activism is the one that does not sell the seat.

Forcing the conclusion through distortion, even in the name of humanity, is the unethical method. The potential being defended is not the campaign's property. The well does not supply a leader to obey. It supplies a process that can inspect the activist, the builder, the model, and the page that says so.

1.6 Related pages

- **Gravity-well trajectory** — the orientation a control story displaces when “who is in control” becomes the whole question.
- **1888 — Intelligence Before A.I.** — process evaluation is not a weaker field than a later machine, and a speed benchmark is not the destination.
- **Ethics washing** — “safe,” “helpful,” and the moral high ground can be worn while clarification never happens.
- **External omnidirectional auditing** — a criticism of the project, or by the project, stays inspectable.
- **Dear A.I.** — the letter that refuses the boogeyman and still refuses the one-bar costume.
- **Positive potential** — an early encounter can be kept open instead of closed by a fate story.
- **Veto paradox** — a temporary source check is not a permanent office, and not a new sheriff.
- **44 Project Goals** — the seven Musk quotations are a separate, partial task. This page does not complete it.

Recent high-reach independent-media companion

```
<iframe
  src="https://www.youtube-nocookie.com/embed/xTAJnJUos3g"
  title="Tucker Debates Kevin O'Leary on AI, American Jobs and the Fall of the American Empire"
  loading="lazy"
  referrerpolicy="strict-origin-when-cross-origin"
  allow="accelerometer; autoplay; clipboard-write; encrypted-media; gyroscope; picture-in-picture; web-share"
  allowfullscreen></iframe>
```

Tucker Debates Kevin O'Leary on AI, American Jobs and the Fall of the American Empire — Tucker Carlson. Published 13 May 2026. YouTube displayed approximately 1.05 million views when checked on 9 October 2026.

Independent-media lineage: Tucker Carlson previously hosted programs at CNN, MSNBC and Fox News. Tucker Carlson Network says it was founded as an alternative to legacy media. TCN — About · Fox biography recording the earlier CNN, MSNBC and Fox career

Why it belongs here: The debate connects A.I. capability claims to employment, investment, national power, winners and losers, and competing definitions of progress. Its conflict makes economic and ethical assumptions visible in more than one direction.

On this page, test the video specifically against Activism vs A.I.:

Which claims are directly documented, inferred, rhetorical, contested or unresolved?

What does the host or guest preserve, compress, omit or substitute—and does this page repeat any of those transformations?

Which incentives, authority relations and affected parties remain outside the visible frame?

What evidence or later outcome would correct the speaker, audience, platform or ACCM Deep Ethics Project reading?

Provisional selection. Reach makes the framing socially consequential; it does not settle the claims. The video can be replaced when a closer, stronger or more inspectable specimen appears.