AI may be to blame for our failure to make contact with alien civilizations

May 11 2024, by Michael Garrett

There's a mindboggling number of planets out there. Credit: NASA/James Webb telescope

Artificial intelligence (AI) has progressed at an astounding pace over the last few years. Some scientists are now looking towards the development of artificial superintelligence (ASI)—a form of AI that would not only surpass human intelligence but would not be bound by the learning speeds of humans.

But what if this milestone isn't just a remarkable achievement? What if it also represents a formidable bottleneck in the development of all civilizations, one so challenging that it thwarts their long-term survival?

This idea is at the heart of a research paper I recently published in *Acta Astronautica*. Could AI be the universe's "great filter"—a threshold so hard to overcome that it prevents most life from evolving into space-faring civilizations?

This is a concept that might explain why the search for extraterrestrial intelligence (SETI) has yet to detect the signatures of advanced technical civilizations elsewhere in the galaxy.

The great filter hypothesis is ultimately a proposed solution to the Fermi Paradox. This questions why, in a universe vast and ancient enough to host billions of potentially habitable planets, we have not detected any signs of alien civilizations. The hypothesis suggests there are insurmountable hurdles in the evolutionary timeline of civilizations that prevent them from developing into space-faring entities.

I believe the emergence of ASI could be such a filter. AI's rapid advancement, potentially leading to ASI, may intersect with a critical phase in a civilization's development—the transition from a single-planet species to a multiplanetary one.

This is where many civilizations could falter, with AI making much more rapid progress than our ability either to control it or sustainably explore and populate our solar system.

The challenge with AI, and specifically ASI, lies in its autonomous, self-amplifying and improving nature. It possesses the potential to enhance its own capabilities at a speed that outpaces our own evolutionary timelines without AI.

The potential for something to go badly wrong is enormous, leading to the downfall of both biological and AI civilizations before they ever get the chance to become multiplanetary. For example, if nations increasingly rely on and cede power to autonomous AI systems that compete against each other, military capabilities could be used to kill and destroy on an unprecedented scale. This could potentially lead to the destruction of our entire civilization, including the AI systems themselves.

In this scenario, I estimate the typical longevity of a technological civilization might be less than 100 years. That's roughly the time between being able to receive and broadcast signals between the stars (1960), and the estimated emergence of ASI (2040) on Earth. This is alarmingly short when set against the cosmic timescale of billions of years.

This estimate, when plugged into optimistic versions of the Drake equation—which attempts to estimate the number of active, communicative extraterrestrial civilizations in the Milky Way—suggests that, at any given time, there are only a handful of intelligent civilizations out there. Moreover, like us, their relatively modest technological activities could make them quite challenging to detect.

Wake-up call

This research is not simply a cautionary tale of potential doom. It serves as a wake-up call for humanity to establish robust regulatory frameworks to guide the development of AI, including military systems.

This is not just about preventing the malevolent use of AI on Earth; it's also about ensuring the evolution of AI aligns with the long-term survival of our species. It suggests we need to put more resources into becoming a multiplanetary society as soon as possible—a goal that has lain dormant since the heady days of the Apollo project, but has lately been reignited by advances made by private companies.

As the historian Yuval Noah Harari noted, nothing in history has prepared us for the impact of introducing non-conscious, super-intelligent entities to our planet. Recently, the implications of autonomous AI decision-making have led to calls from prominent leaders in the field for a moratorium on the development of AI, until a responsible form of control and regulation can be introduced.

But even if every country agreed to abide by strict rules and regulation, rogue organizations will be difficult to rein in.

The integration of autonomous AI in military defense systems has to be an area of particular concern. There is already evidence that humans will voluntarily relinquish significant power to increasingly capable systems, because they can carry out useful tasks much more rapidly and effectively without human intervention. Governments are therefore reluctant to regulate in this area given the strategic advantages AI offers, as has been recently and devastatingly demonstrated in Gaza.

This means we already edge dangerously close to a precipice where autonomous weapons operate beyond ethical boundaries and sidestep international law. In such a world, surrendering power to AI systems in order to gain a tactical advantage could inadvertently set off a chain of

rapidly escalating, highly destructive events. In the blink of an eye, the collective intelligence of our planet could be obliterated.

Humanity is at a crucial point in its technological trajectory. Our actions now could determine whether we become an enduring interstellar civilization, or succumb to the challenges posed by our own creations.

Using SETI as a lens through which we can examine our future development adds a new dimension to the discussion on the future of AI. It is up to all of us to ensure that when we reach for the stars, we do so not as a cautionary tale for other civilizations, but as a beacon of hope—a species that learned to thrive alongside AI.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation