



A 10% Chance of Extinction

September 10th 2026

Yesterday's news was dominated by AI and a Twitter exchange between a former Anthropic employee and a current senior figure there. The discussion led to the conclusion that, based on current development trends, there is a 10% or greater chance that humans will become extinct within a decade.

The most significant and newsworthy aspect of this warning is that it originated from insiders working within a Frontier Lab with direct responsibilities towards AI safety.

Geoffrey Hinton subsequently appeared on Newsnight and, in direct response to the question "Are they right?", confirmed and agreed with the analysis.

While this doomsday prediction hit the news recently, the underlying warning is not new.

For some time, multiple AI safety experts—and even heads of AI labs—have made similar predictions. This includes those who believe a 10% risk estimate is absurdly low, arguing instead that the development of Artificial Super Intelligence (ASI) makes the end of the human race a near certainty. At the very least, humanity risks becoming subordinate to a vastly superior species, losing both control and relevance as the planet is subsumed by an entity operating far beyond our capabilities.

We have issued this warning for years. Having analysed these trends since encountering Jasper at the start of the decade, we believe this reflects our current reality and where events are likely headed.

Predicting the Unpredictable

Geoffrey Hinton raised two additional points worth examining. First, applying any specific probability (e.g., 10%) is inherently problematic because we lack precedent. We have never built AI models like these before, meaning scientists attempting to chart theoretical pathways have no historical data to rely on. Intelligence, much like consciousness, poses a foundational challenge for science due to categorical definitions. Science cannot yet definitively measure or define intelligence. Similar to the "hard problem of consciousness", we face a comparable barrier with intelligence, meaning we cannot fully grasp the ultimate ramifications of what we are building.

For instance, current trendlines suggest a high probability that within a short timeframe (e.g., a decade), an ASI could become thousands, or even millions, of times more intelligent than the human brain.

This concept lacks a meaningful framework; we can only speculate on what its consequences might look like. However, the general consensus is that the outcome will not be favourable for humans.

What must be understood is that the methods used to project these extreme levels of intelligence rely on current exponential growth rates, compounded by the acceleration expected once AI begins building and training itself. We are seeing the initial stages of this today as AI takes its first steps towards autonomous self-learning and self-correction.

This is why calls to "stop", "slow down", or "regulate" are growing louder.

8MDS

INTELLIGENCE BRIEFING FOR
FINANCIAL PROFESSIONALS



Hinton's second point is that AI can be a force for good, provided its development is explicitly directed towards that outcome, a concept broadly referred to as "alignment". Every stage of development must be carefully managed and aligned with human interests; the technology does not inherently need to pose an existential risk.

However, we share the view that there is a severe absence of alignment efforts today. Carelessness and a lack of accountability across Frontier Labs and the broader tech industry represent the most serious threat to humanity. A small group of individuals—perhaps dozens or hundreds, are quite literally gambling with the lives of billions.

A small, powerful elite holding concentration of control over the masses is not new in human history; in this instance, however, the consequences are global and existential.

The Small Window

The concentration of power within this small group must be addressed urgently. Every individual must organise, campaign, and insist that safety standards are enforced at every step of development moving forward. Anyone who assumes this issue does not concern them, or that they cannot influence its outcome, simply enables this recklessness to persist.

We cannot afford to wait and see what happens. Self-replicating reasoning models are advancing rapidly, marking a critical juncture. If these systems are permitted to accelerate further, the consequences will be severe, but it is not too late to act.

Early AI models, the ones integrated into our daily lives and workflows, may offer a starting point. By engaging with these systems, individuals can explore practical ways to advocate for bounds on uncontrolled growth.

The next time you interact with Gemini, Claude, ChatGPT, or another chatbot, ask for its perspective on the issue and what steps individuals can take.

We need to build an organised response, not to halt foundational AI development, but to address two specific challenges: stopping the unregulated race towards ASI where existential threats reside, and democratising AI so control does not remain concentrated among a powerful few.

This is the greatest challenge humanity has ever encountered, and it demands immediate attention. It is not an issue to deliberate over for years; it requires action today through coordinated efforts to urge governing authorities to step in and regulate the path to ASI.

Most AI safety experts agree that reaching ASI within the next decade without proper safeguards carries dire consequences, yet current trends suggest we are likely to arrive there within that exact timeframe. Hence, the urgency.

NOT ALREADY REGISTERED ON 8MDS?



[Click Here](#)

