SAFE AI FOUNDATION

5 Upcoming Potential Dangers of AI

Chai K Toh

Safe AI Foundation

19 JUNE 2026

1 Dangers of Recursive AI

AI recursively improving itself without human control or oversight can be a recipe for danger, risks and out of control. AI recursively learning on its own (via the internet, for example) or feeding data from sensors (gossips, human conversations, etc.,) or from other AIs can spiral into an uncontrollable state. Given the immense autonomy and the ability to acquire an immense amount of knowledge, the AI can run the risks of:

1. Not knowing WHEN TO STOP
2. Not knowing it has reached a dangerous level of intelligence
3. Not knowing it has learnt a lot of bad things
4. Totally bypassing any human means to stop its operation
5. Becoming too clever, too sneaky, replicating, hiding, etc.
6. Ganging up with other bad AIs
7. etc.

Lack of Limits & Human Supervision: AI recursively improving itself without human intervention and oversight is considered a high risk. AI learning from other AIs recursively is also regarded as dangerous. The issues lie in the lack of limits and no human oversight and supervision on what AI is learning. It can be consistently learning bad or wrong things. Furthermore, this recursive learning does not stop, making it uncontrollable. **Uncontrollable AI should not be released to the public.**

2 Dangers of AGI & ASI

Momentum – A group of engineers and VCs have been pursuing to realize AGI. There is no solicitation for public feedback, government's consent, and founders have pursued to work on this for various reasons, be it: (a) cool, (b) eye and ears catching, (c) lure VCs for investments, etc.

Why the Need for AGI? – It is still unknown till today why we even need AGI.

- For what?
- What is it good for?
- Can we live without it?
- Can it boost up our economy, our human capital, our job numbers, etc?

Lack of Clarity: There hasn't been a detailed investigation and study on AGI, released by an authority. Worse still, there isn't an agreed definition of AGI and ASI, what they can and cannot do.

The Dangers: If AGI and ASI way surpass human intelligence and are fully autonomous and highly distributed, humans can lose control of it and AI can pose a threat. Self-governing and self-improving AGI and ASI that outsmart humans mean we will have a hard time stopping them.

3 ▒ Dangers of Over-Scaling

The Possible Misconception: There has been a constant push to scale up, buy more servers, GPUs, etc., and host more data centers. Somehow, somewhere, news spread that we need to scale up to make AI more intelligent and run faster. However, **there is no truth out there to confirm that scaling compute resources will make AI more intelligent**. There is no universal agreement on this point. Hence, scaling does not directly equate to AI intelligence. Scaling compute does make things run faster, but not necessarily more intelligent.

Energy Crisis: Secondly, while scaling speeds up execution and reduce latency, it does result in huge high energy consumption and large investments needed to build data centers, requiring both LAND and SERVERS (chips, networks, cooling systems, etc.,) This imposes considerable investment and spending on hardware resources, land, and energy and is not entirely environment friendly. The carbon footprint created will be significant.

The Question: If scaling does not directly equate to AI intelligence, then what is the point of spending billions to scale up data centers? Shouldn't investments be spent on **software** - developing better **AI algorithms and architectures** so that they are more intelligent, accurate, efficient, and safe to use?

4 ▒ Dangers of Uncontrollable AI with Total Autonomy

We are currently at Stage 1 of our AI progression roadmap, where we are developing AI technologies and concurrently deploying some of the AI products completed. Hence, we have gradually moved from basic AI research to AI product development and deployment.

Issue of Controllability: GPTs (LLMs) are followed by AI Agents, "who" can now work autonomously, without human intervention. While AI companies claimed that AI Agents are controllable software entities, several incidents have revealed otherwise. (Source: please google)

Dangers of Uncontrollable AI: In the SAFE AI Foundation's response to the White House OSTP's Call For Comments, we explicitly stated that **uncontrollable and fully autonomous AI should not be released to the public**, as it can pose tremendous threat to public safety. Once such agents are deployed, it can be hard to trace and stop. With privileged access rights to systems and IT, malicious agents can cause unrecoverable damages.

Need For Regulations: Any creations of such uncontrollable entities can result in public and enterprise harm, once deployed. New regulations are needed to hold creators of such malicious agents accountable.

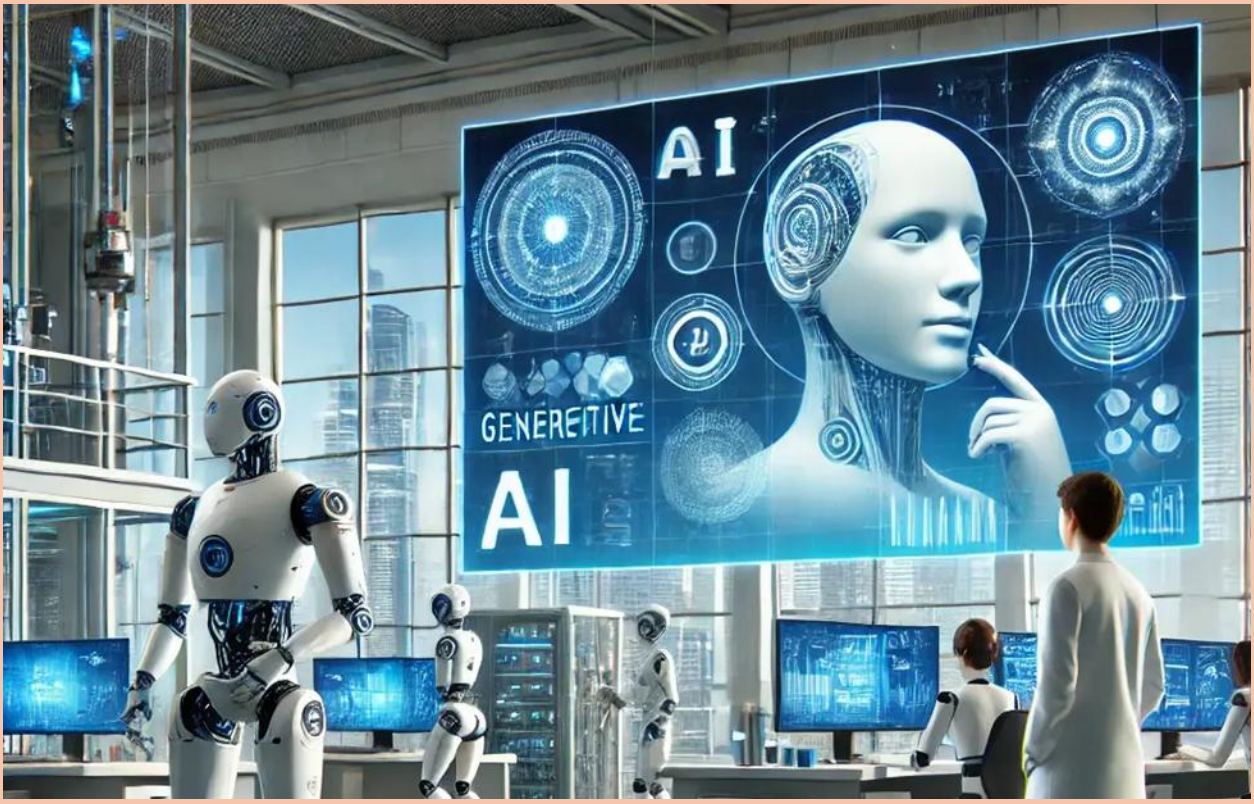
Our view on this still stands today.

AI Impact Roadmap: see: https://safeaifoundation.com/ai-impact-roadmap

5 ▒ Dangers of Misuse, Abuse, Weaponization & Sole Autocracy

While uncontrollable and malicious AI Agents pose dangers, the misuse and abuse of such entities by humans are equally liable. Persons or hostile nations with such AI powers can or may pose a threat to others.

1. Negative Side Effects:
    a. Cognitive Offload
    b. Human Intelligence Hijack
    c. Human Addiction to AI
    d. Humans' Loss of Creativity
    e. Humans' Loss of Originality
2. Misuse
    a. Various Types of Misuses
    b. Deepfakes, Fake News, Fraud, Copyright Violations, etc.
    c. Punishments for Misuse
3. Abuse
    a. Misuse of AI several times = Abuse
    b. Punishments for Abuse
4. Weaponization
    a. Terrorists group can misuse AI to attack people or nations
    b. Terrorists or cults may harness AI to make weapons
5. Sole Holder of AI Supreme Power (Autocracy)
    a. The most intelligent and deadly AI Model can "threaten" and "rule" the world... (mentioned by many sources)
    b. It can be used as a bargaining chip
    c. It can over-rule other competitors
    d. It can gain superiority and command conformance



SAFE
AI
FOUNDATION

www.safeaifoundation.com

The SAFE AI Foundation was formed in 2025, in the heart of the Silicon Valley, California, USA. Its mission is to embrace the AI revolution and promote the use of good AI to improve on the quality of life and the economy. It will address the issues of bad AI and its impacts on society. The foundation will inform, educate, and mitigate the dangers of bad AI and help the society to embrace good AI. The foundation will interact with industries, universities and public sector to jointly advance the progress of Good AI. The foundation is a non-profit 501c(3) organization, registered in the State of California, USA.