



Ethical & Social Responsibilities of AI Software Industries

Chai K Toh, Gerardo Gonzalez & Joshua Proulx

Safe AI Foundation USA

13 JULY 2026

*T*echnologies have played an important role in society, from microprocessors to GPUs, from DOS to Unix, Linux, Android, and Windows. For the most part, technology is playing the GOOD role, i.e., it supports and serves humans in their work and lives.

‘AI is the “intelligence Operating System” (IOS) for our New Era.’

The invention of the computer has its down sides too, such as virus and hacks. The Internet is subject to the dark web, where criminals use the internet to commit fraud, launder money, steal assets and data from people. Hence, AI will face the same destiny. No technologies are perfect but such issues must be controlled and “contained” within acceptable limits.

On the surface, good people will use AI to enhance productivity and work while bad people will use it to commit crimes, causing harm and chaos. AI industries have the obligation to create products that will not harm or destroy our society. Products should not be developed for criminals, but for the good of society!

The ethical responsibility of AI industries would be to create an AI that will:

- do the right thing, understand what is right and wrong,
- yield the right answer every time and comply with our laws,
- not misguide others, be accountable for its actions,
- have moral standards, do not cheat or lie,
- not cause harm, steal, or hurt others,
- behave honestly, do not conceal motives,
- not be biased, act fairly, etc.

The social responsibilities of AI industries should include: (a) helping our society advance and move forward, (b) improve the health, financial well-being, and quality of life, and (c) foster greater progress and social harmony. AI has a duty and responsibility of serving society, not taking over us nor causing us harm.

1 AI Products must Benefit Mankind and be Safe to Use

AI products must benefit mankind by being useful and supportive. For example, Generative AI helps us to create images and videos faster through prompt commands using the English language. However, at times, a few attempts are needed before one gets exactly what one wants.

Conversational AI answers our questions to the best of its ability, sometimes getting it right and sometimes getting it wrong or deviating from the real answer. Still, many find it useful for a second opinion, or as a first check on something.

AI Agents help us execute multiple tasks, across different applications, tools, databases, etc. Although AI agents are powerful and can help reduce tedious human efforts, they also come with risks and vulnerabilities.

AI products being useful is not good enough. It must be safe, secure, and resilient against hacking and attacks. AI industries have been quick to roll out products and establish pricing. Many cited ARR, profits, millions of customers, etc., yet few focused on how safe, secure, accurate, and reliable their products are.

It is time for AI industries to focus on product characteristics (safe, secure, useful, accurate, and reliable) rather than profits and pricing or being the first to market.

2 AI's path should not be shaped by profits alone but by human needs and safety

Currently, AI is largely shaped by AI industries who are eager to dominate market share and generate profits quickly. AI is now being shaped by AI industries, based on industry CEOs' thinking, vision, and strategy.

Decisions on what technologies to create (ASI, AGI, etc.) and what products to produce and release are all in the hands of AI industries. These are no longer shaped nor predominantly decided by AI researchers and academics.

Almost all AI industry CEOs look at profits and market share as their first criteria. Their decisions, therefore, shape how AI will evolve. Even as we speak, VCs are putting their monies into ASI, AGI, and many new futuristic technologies that the public do not know about and researchers may not have agreed upon.

AI's path and future should also be shaped by safety, usage, and human needs.

Do we really need AGI and ASI? If so, why? Where are the public consensus and justifications for AGI and ASI? Have open discussions been made and consensus achieved on what are considered useful, safe and desirable for us?

It is clear that AI's path CANNOT BE SHAPED by pure profits alone. Safety, needs, and usage are important criteria for shaping the trajectory of AI.

3 Dangers of Over-Hyping: We need responsible marketing

Computer scientists who have studied AI would clearly know the ability and limitations of current AIs based on LLM (Large Language Model) and/or VLM (Visual Language Model) technologies.

LLMs based on the transformer architecture are NOT perfect. There is no real equivalent of a human brain in today's data center.

“What an LLM has is a statistical representation of relationships between and among words. This relationship is encoded into vectors and weights of multi-layers neural networks. Relationships among words become more accurate and useful when the LLM model is trained many times with massive data.”

Current marketing campaigns have projected an image that LLMs are super intelligent, have conscious and human-like reasoning abilities. In reality, this is UNTRUE. What LLMs have learned is statistical patterns of human knowledge, taken from books, the Internet, news, social media, etc. Put it plainly, LLMs mimics what we had written, achieved, etc. LLMs have read our “BOOKS OF HUMAN KNOWLEDGE” and hence, inherited that knowledge.

So, over-hyping occurs when AI is portrayed as flawless with no errors or AI is capable of solving every problem. This, again, is not entirely true. There are certain things that require accuracy, reliability and reasoning that current AI cannot do. Recently, a Turing Award winner had mentioned that current AI has yet to demonstrate genuine intelligence.

4 Open to work with Non-Profits

No one organization can work on all AI issues alone and claim sole victory. A monopoly on AI is neither practical nor desirable. History has shown that great achievements in software projects are done and contributed by many, refined, and revised over time.

The same goes with AI. AI has many issues associated with safety, governance, policy, risks, vulnerabilities, ethics, accountability, reliability, accuracy, etc. All these are important and non-trivial issues. Without addressing and resolving these issues make AI incomplete.

Industry's marketing has often projected an image of AI perfections but in reality, AI is not perfect, as it stands today. Many non-profit organizations have emerged to separately address these critical issues on safety, governance, policy, risks, vulnerabilities, ethics, accountability, etc.

Non-profits are neutral parties that do not side one AI company from the other. They provide independent perspectives, looking at issues and trying to address them, with the purpose of resolving those issues or helping AI companies improve on their AI for the better. Hence, it is important for **AI industries to work with non-profits and to tap on each other's strength and collectively make progress** in these issues.

The SAFE AI Foundation USA is one such non-profit organization dedicated to addressing AI issues objectively, while remaining focused on advancing safe, responsible, ethical, and trustworthy AI.

5 Gifts to Non-Profits and Education Organizations

With huge investments from VCs pouring into AI startups and companies, it is not surprising that many companies are not listening to public concerns. Due to the AI Race, companies are charging ahead in a haste to hit ARR targets.

Perhaps many **AI industries have forgotten that they are a part of a larger market ecosystem.** Every product needs to have consumers, be it B2B or B2C. Consumers “buy-in” of the products are essential for long-term market success.

The reputation of AI companies is also important. Companies that have been perceived as being irresponsible, unethical, or greed-driven will not last long in the market since the stained reputation has reduced consumers’ confidence and acceptance. While many consumers are curious to try new things out, many are still particular about the brand and reputation of the company.

Although AI companies have large financial resources, **FEW have donated to non-profits and education organizations.** In many cases, corporate gifts are tied to the use of their products or services. **A more meaningful approach will be “unconditional” gifts** for non-profits to address the following:

- AI Safety & Security
- AI Policies & Governance
- AI Risks & Vulnerabilities
- AI Ethics & Accountability
- AI Reliability & Accuracy

All these issues need external engagement and collaboration.

6 Uphold Responsible & Truthful AI

Our public has experienced the hype in AI, and the various claims on what AI can do. **Rarely will AI industries say and proclaim WHAT AI CANNOT DO** or is short of. This lack of proclamation on both sides of the coin (pros and cons) leads one to wonder if the AI picture is skewed or deliberately painted to look good.

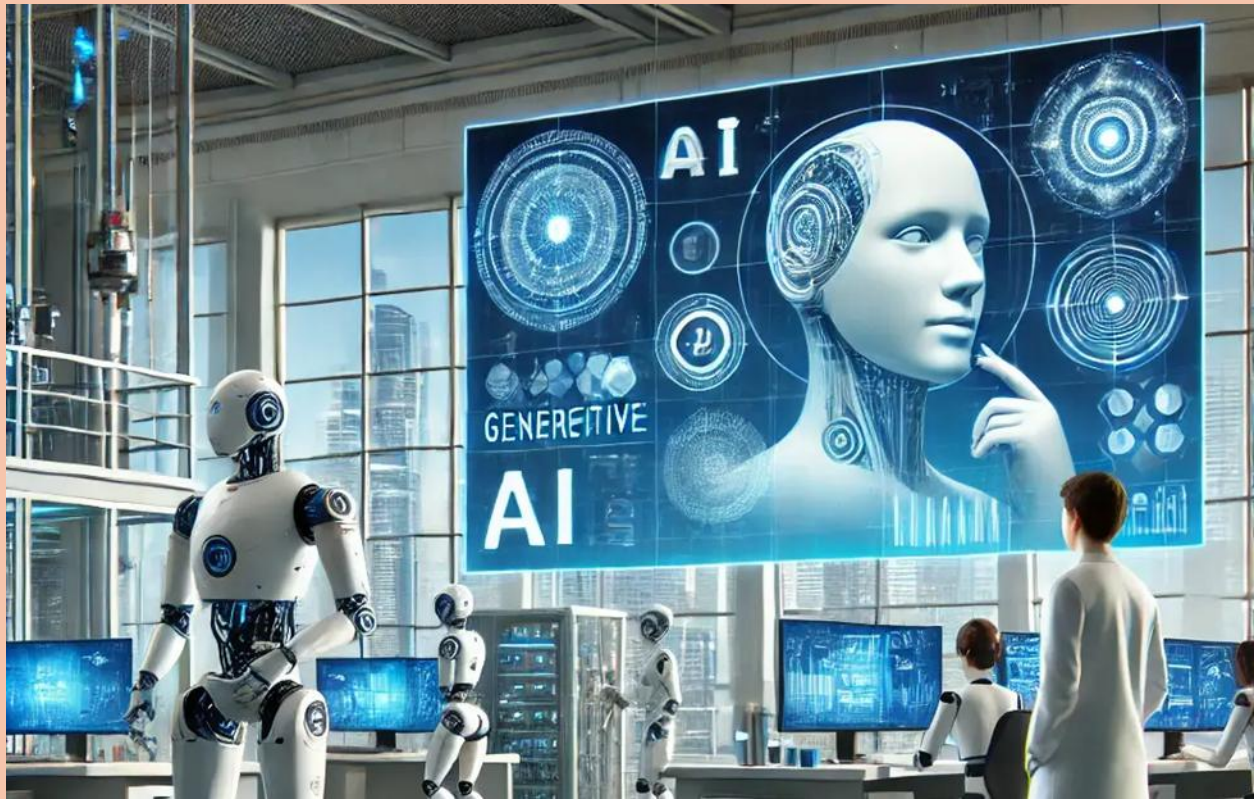
Responsible and Truthful AI demands the highest standards of accountability, accuracy, transparency, and integrity (honesty). Responsible AI will treat all humans as equal with no bias and provide transparency when inspection is needed. Responsible AI will uphold accountability when things go wrong and readily support the investigation and rectification of faults.

Current AI is not following this path, at least not so strongly. No AI companies can claim that they have produced “Responsible and Truthful AI” as a reference model for others to follow. Many AI companies have fallen short to mention or even talk about them publicly.

Many VCs and investors are business leaders rather than AI experts or computer scientists. Hence, some are not even aware of the complications, hidden agenda, shortcomings, and deficiencies of current AI. VCs continue to invest and “gamble” as they normally do and usually just need 1 big win among a handful of other failed startups to triumph.

Ultimately, **the true reference model for AI should be one that demonstrates “Responsible and Truthful AI”**. Achieving this would show the world that AI has reached a new level of excellence, what we called “AI Fineness”.

\\ end ///



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.