

---

## Five Reasons & Two Recommendations for AI to be a National Security Priority Globally

---

“Big things have small beginnings.”

— [Said](#) by the android David before deliberately contaminating a smug human scientist’s drink with dark genetic fluid. *Prometheus* (2012)

### The Argument

This paper argues that if AI becomes a national security priority globally, states will more easily understand why and how they must establish an enforceable regulatory regime that prevents catastrophic AI risks. The paper offers five reasons why private companies, nonprofits and philosophers appear to be the fittest today to enable rapid, actionable and policy-relevant sense-making of the unprecedented risks of AI; and two practical recommendations to contribute to the implementation of this sense-making.

AI has been in the global social discourse since decades, but only recently is it gaining currency in the West thanks to the—nearly entirely private—R&D shift towards “a previously fringe approach (neural networks) [that] has become mainstream and highly” lucrative<sup>i</sup>. In its recent paper “Natural Selection Favors AI over Humans,” the nonprofit Center for AI Safety states “at some point ... AIs ... could outcompete humans, and be what survives<sup>ii</sup>,” because “... AIs will likely be selfish” and smarter than humans; thus posing “catastrophic risk” to humanity.

The first indicator of the plausibility of this risk is that authoritative AI industry insiders are warning the world that AI could pose an existential threat. Second, with the exception of the USA and China, governments are irreversibly losing to private companies the race to procure the computing and human resources needed to research, develop and roll out safe and reliable AI technologies<sup>iii</sup>. Third, the most authoritative industry insiders are unaware of the technology’s full range of capabilities. Fourth, despite this, the AI industry is developing exponentially within a global regulatory void. Lastly, if AI could redefine Society and even supplant human civilization, philosophy reemerges as fundamental to policy.

### ***If the rise of transformative AI is far more political than technological, then policy is the key function***

Truly enforceable AI safety and security for humans requires political awareness as well as both political and societal consensus; which would ensue if sovereign States would deem AI a top national security priority. This Theory of Change finds support in Anthropic’s Policy Director’s claim that now is “more a political moment than a technological one,”<sup>iv</sup> because of the unprecedented societal implications of having this hugely transformative, lucrative and unregulated technology mostly in the hands of private companies<sup>v</sup>.

Starting from the premise “AI safety is highly neglected<sup>vi</sup>,” two recommendations seem immediately fit to support ongoing efforts by some private companies and nonprofits to research, document and divulge key evidence about this fact; and leverage growing consensus in the industry about concerns over AI while extending and globalizing the said consensus to enable permanent human-centric and enforceable AI safety.

The first recommendation is to craft, with private and nonprofit leadership, a global coalition for safe AI with a whole of society / bottom-up, and a whole of nations / top-down approach; engaging key persons and organizations in a Global AI awareness network from which to source knowledge and ideas for the normative rules to ensure that ethics and integrity drive AI always. Secondly, to work with governments for a robust regime of global safeguards based on experiences in nuclear, counter-terrorism, and drug control.

### **Conclusion: *Love Thy Human Cognition***

Never subordinating human cognition to a non-human conscience seems a prudent option to keep humanity safe from the risk of losing control to conscious machines. The reemergence of philosophy as critical science to mitigate the potentiality of this existential risk, attests to the perennial value of ancient human myths and archetypes that still render the human experience intelligible, and thus, more predictable.

Nothing in this world is more powerful than love. Love for Humanity—for us the chiefest of creations in the Universe — should guide AI development always, within a regulatory framework with coercion and safety and security standards drawn from other fields essential to national security.

## Endnotes

- i "Securing Liberal Democratic Control of AGI through UK Leadership," James W. Phillips, 14 March 2023. <https://jameswphillips.substack.com/p/securing-liberal-democratic-control>
- ii Dan Hendrycks, "Natural Selection Favors AIs over Humans," arXiv:2303.16200, March 2023. <https://arxiv.org/abs/2303.16200>
- iii Phillips, "Securing Liberal Democratic Control of AGI through UK Leadership" (see n. 1).
- iv Jack Clark, "Import AI 321: Open Source GPT3; Giving Away Democracy to AGI Companies; GPT-4 is a Political Artifact," *Import AI*, 20 March 2023. <https://importai.substack.com/p/import-ai-321-open-source-gpt3-giving>
- v Clark, "Import AI 321" (see n. 4).
- vi "About," Center for AI Safety. <https://www.safe.ai/about>