

# Former Heads of State, Nobel Laureates, and Thought Leaders Urge International Agreement on AI “Red Lines”

**September 22, 2025**

Launched during the 80th session of the UN General Assembly, the Global Call for AI Red Lines urges governments to reach an international agreement on “do-not-cross” limits for artificial intelligence by the end of 2026 to prevent the most severe risks to humanity and global stability.

The campaign’s unprecedented range of signatories signals a new phase of global concern, expanding beyond the AI scientific community to leaders in economics, human rights, and diplomacy.

The Global Call for AI Red Lines warns that without binding international rules, humanity faces escalating risks—from engineered pandemics and large-scale disinformation to global security threats, systematic human rights abuses, and loss of human control and oversight over advanced systems

Drawing on precedents set by past international agreements addressing global challenges such as biological weapons and human cloning, the campaign calls upon governments to reach agreement to prohibit specific AI uses and behaviors deemed universally too dangerous or high-risk, and back those prohibitions with enforcement.

Signatories include more than 200 prominent individuals and 70 organizations worldwide, such as:

- 9 former heads of state or ministers, such as Mary Robinson, Juan Manuel Santos, and Enrico Letta;
- 10 Nobel laureates, including Joseph Stiglitz, Maria Ressa, Daron Acemoglu, and Geoffrey Hinton;
- AI pioneers and Turing Award winners, including Yoshua Bengio and Andrew Yao;
- Policymakers, including Audrey Tang, Taiwan's first Digital Minister, and Csaba Kőrösi, former President of the UN General Assembly;
- Leading public intellectuals, such as Yuval Noah Harari and Stephen Fry;
- Industry leaders and executives, including OpenAI co-founder Wojciech Zaremba, Google DeepMind’s Ian Goodfellow, and Anthropic’s CISO Jason Clinton.

The initiative is co-organized by the French Center for AI Safety (CeSIA), The Future Society, and UC Berkeley’s Center for Human-Compatible Artificial Intelligence, alongside 20 partner organizations.

## **Present Dangers, Urgent Warnings**

The launch of the Global Call comes amid growing concern around AI incidents. Recent tragic cases of AI-driven delusion and suicide are only early glimpses of far greater dangers that lie ahead. Hundreds of millions of people already rely on advanced AI systems, and leading AI developers indicate their latest models are approaching dangerous capability thresholds. The Global Call stresses that governments must act before the window for meaningful intervention closes.

Signatories emphasized the stakes:

**Yoshua Bengio**, 2018 Turing Award Laureate:

*"The current race towards ever more capable and autonomous AI systems poses major risks to our societies and we urgently need international collaboration to address them. Establishing red lines is a crucial step towards preventing unacceptable AI risks."*

**Maria Ressa**, Nobel Peace Prize Laureate:

*"Without AI safeguards, we may soon face epistemic chaos, engineered pandemics, and systematic human rights violations. History teaches us that when confronted with irreversible, borderless threats, cooperation is the only rational way to pursue national interests."*

**Stuart Russell**, Distinguished Professor of Computer Science at the University of California, Berkeley:

*"The development of highly capable AI could be the most significant event in human history. It is imperative that world powers act decisively to ensure it is not the last."*

**Ahmet Üzümcü**, Former Director General of the Organization for the Prohibition of Chemical Weapons (OPCW), which was awarded the 2013 Nobel Peace Prize:

*"It is in our vital common interest to prevent AI from inflicting serious and potentially irreversible damages to humanity, and we should act accordingly."*

**Yuval Noah Harari**, Author and Professor of History:

*"For thousands of years, humans have learned—sometimes the hard way—that powerful technologies can have dangerous as well as beneficial consequences. With AI, we may not get a chance to learn from our mistakes, because AI is the first technology that can make decisions by itself, invent new ideas by itself, and escape our control. Humans must agree on clear red lines for AI before the technology reshapes society beyond our understanding and destroys the foundations of our humanity."*

**Csaba Kőrösi**, Former President of the UN General Assembly:

*"Humanity in its long history has never met intelligence higher than ours. Within a few years, we will. But we are far from being prepared for it in terms of regulations, safeguards, and governance."*

## **Coordinated Action Needed Today**

While some governments have begun to develop their own AI policies and regulations, the Global Call stresses that fragmented national efforts will not be sufficient to manage a borderless technology. Only an international agreement can ensure that AI is developed and deployed safely, securely, and in line with shared fundamental freedoms.

The Call [was announced](#) this morning by Maria Ressa at the high-level UN meeting marking the organization's 80th anniversary, and will be presented at the AI Safety Connect event at UN Headquarters in New York on Thursday, September 25, 2025.

**The full statement and list of signatories and partners are available at: [red-lines.ai](https://red-lines.ai)**

---

**The French Center for AI Safety (CeSIA)** is an independent non-profit based in Paris dedicated to preventing major AI risks. It provides expert policy advice and conducts technical research to help institutions respond to the urgent need for AI governance. CeSIA aims to help establish safety guardrails for the safe development of advanced AI.

[www.securite-ia.fr](https://www.securite-ia.fr)

**The Future Society** is an independent non-profit organization founded in 2014, with a mission to align artificial intelligence through better governance. With a team of AI experts based around the world, The Future Society regularly advises U.S., EU, and global policymakers on how to address the threats of today and tomorrow.

[www.thefuturesociety.org](https://www.thefuturesociety.org).

**The Center for Human-Compatible AI (CHAI)** is a research center at UC Berkeley focused on ensuring that powerful AI systems are beneficial to humanity. CHAI advances technical research on value alignment and decision theory, building the foundations for AI systems that are provably aligned with human preferences.

[humancompatible.ai](https://humancompatible.ai)

Contact: [media@red-lines.ai](mailto:media@red-lines.ai)