

Launch of the report on global compact on extreme AI risks

Strategic Foresight Group

**Opening remarks by
Ambassador Thomas Greminger, Executive Director, GCSP**

Geneva, 6 July 2026



Ladies and Gentlemen,

Welcome to the launch of this report on the Global Compact on Extreme AI Risks. I would like to thank the IFRC and the ICT4Peace Foundation for their cooperation. Most of all, I would like to congratulate my friend Sundeep Waslekar and his team at the Strategic Foresight Group for this report.

I think we all share a common concern that artificial intelligence is developing much faster than our ability to regulate it.

This is not just a challenge of bringing a new form of technology under control. It poses existential risks for the future of humanity.

Scientists – including pioneers of so-called “deep learning” – have warned that we may lose control of autonomous AI systems. The most acute risks are in domains such as cybersecurity and biotechnology, as well as disinformation. There is also an increased risk of military escalation and even unintentional nuclear war. Three plausible – and frightening – scenarios are outlined in the report.

Some of these risks may arise from intentional misuse of AI. Others could be unintended consequences caused by a loss of control or misalignment with human intent. Either way, the consequences could be devastating – even irreversible.

Indeed, in May 2023, a group of scientists and eminent people – under the auspices of the Center for AI Safety – issued a short statement saying that: “Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.”

Thus far, most of us don’t see or even think about such possibilities. But AI is evolving so quickly that we must be aware of the dangers – and prevent them before it is too late.

Ladies and gentlemen,

I would like to stress that my speech was written by a human being.

Perhaps, in a few years, we won’t need speechwriters anymore: machines will do the thinking and writing for us.

But what happens when we take humans out of the loop for other decisions?

There are, for example, concerns about lethal automated weapon systems where humans are no longer in the chain of command. Or overreliance on big data when making rapid decisions on choosing military targets.

Pope Leo the 14th rang the alarm bell in his recent thought-provoking Encyclical “On Safeguarding the Human Person in the Time of Artificial Intelligence”.

As he pointed out, machines do not have a moral conscience, since (as he said) “they do not judge good and evil, grasp the ultimate meaning of situations, or bear responsibility for consequences.” They cannot be held accountable.

Agentic AI could therefore increase the risk of dehumanizing conflict – rendering it more impersonal, lowering the threshold for resorting to violence, and reducing victims to data. He goes so far as to say that AI should be “disarmed”. By that he means “freeing it from the mentality of ‘armed’ competition and a race for ever more powerful algorithms and larger datasets, driven by the desire to secure geopolitical or commercial dominance.”

How can this be done? As the Pope admits, it is not enough to make moral arguments or invoke ethics. He calls for “robust legal frameworks, independent oversight, informed users and a political system that does not abdicate its responsibility.”



Thus far, as pointed out in the report being launched today, responses to regulating AI have been national, reactive and disjointed. Greater cooperation is required, as is a more global perspective.

There are some encouraging signs in this direction.

Two years ago (in May 2023), the Elders issued a statement calling on world leaders to “work together urgently on the design of strong international governance, to allow all humanity to take advantage of the opportunities of AI, while limiting the enormous risks.”

This was followed up by The Bletchley Declaration made at the AI Safety Summit of November 2023. As it pointed out: “Many risks arising from AI are inherently international in nature, and so are best addressed through international cooperation.” – not only between states, but also involving civil society, academia and especially companies.

The OECD has called for anticipatory technology governance – this is vital since AI is being developed at speed and scale, while diplomacy is lagging behind.

The UN General Assembly has established an Independent Scientific Panel on AI and a Global Dialogue on AI Governance. And the Security Council has looked into the implications of AI on peace and security.

Furthermore, the Netherlands and South Korea have spearheaded a series of worldwide consultations on Responsible AI in the Military domain. These meetings have resulted in declarations endorsing the centrality of humans in critical military technologies.

So there is some momentum.

The key will be to get enough influential stakeholders pulling in the same direction. China and the US – which are two AI superpowers – are apparently exploring guardrails, particularly preventing malign actors from accessing the most powerful or advanced AI models. Other states, and certainly civil society, should also realize the self-interest in greater regulation.

Sceptics might argue that the current fractured geopolitical situation is not the most conducive environment for brokering new international agreements. But in the past two years, states – even with fundamentally different world views – have agreed to a UN Convention against Cybercrime and a Global Ocean Conservation Treaty. So it’s possible.

Furthermore, as pointed out in the report being launched today, the bar for cooperation can be rather low. Shared threat perceptions may be sufficient to enable international cooperation, even among strategic rivals.

Convergence in threat perceptions can lead to a convergence of responses. This paper – entitled The Essential Convergence – calls for a coordinated framework through which states align around shared definitions of extreme risks, minimum safeguards, and collective response mechanisms. This is a rather light form of regulation that can create interoperable global mechanisms capable of managing extreme risks collectively.

That sounds like common sense to me. And highly urgent.

I therefore look forward to hearing more from Sundeeep about his proposal for a Global Compact on Extreme AI Risk.