Parliamentary action on AI webinar series

Extreme risks from AI: The parliamentary role in setting limits

22 October 2026

Key takeaways

On 22 October 2026, the Inter-Parliamentary Union (IPU) hosted a webinar on the theme Extreme risks from AI: The parliamentary role in setting limits, organized in cooperation with the United Nations Development Programme (UNDP), UNESCO, the Commonwealth Parliamentary Association and the Internet Governance Forum. It was the third webinar in the *Parliamentary action on AI* series, which examines how parliaments are addressing the impact of artificial intelligence (AI) through their core functions, and aims to facilitate exchanges and peer learning between parliamentarians and staff working on similar issues.

The webinar addressed the parliamentary response to extreme risks from AI, which include the use of the technology to develop biological and chemical weapons, as well as its integration into nuclear deterrence, the increasing autonomy of AI systems and the facilitation of cyberattacks. What these risks have in common is that they could pose existential threats to humanity.

In response to this urgent challenge, the IPU brought together parliamentary perspectives from Canada, Costa Rica, Romania, the United States of America and the European Parliament, alongside experts from the Center for AI and Digital Policy (CAIDP) and AI Safety Connect (AISC). The webinar aimed to foster inter-parliamentary exchange and identify concrete steps parliaments can take to address extreme AI risks.

The panellists agreed that extreme AI risks are no longer abstract threats but, increasingly, a source of concrete harm. Many jurisdictions have begun moving from voluntary commitments towards binding rules on transparency, incident reporting, independent testing and emergency powers. Establishing common minimum standards for frontier AI has emerged as a critical priority for parliaments. Recognizing that no parliament can address these risks alone, the speakers called on parliaments to legislate proactively, engage with experts and strengthen international coordination.

Resources:

- Video of the webinar on the IPU's YouTube channel on artificial intelligence
- IPU issue brief: *Democratic control of artificial intelligence: The role of parliaments in addressing extreme risks* (October 2026)
- IPU op-ed: "AI urgently requires democratic leadership" (September 2026)
- Resource sheet on extreme AI risks
- IPU resolution: *The impact of artificial intelligence on democracy, human rights and the rule of law* (October 2024)
- IPU resolution: *Addressing the social and humanitarian impact of autonomous weapon systems and artificial intelligence* (March 2024)
- Kuala Lumpur Declaration on *Parliaments and responsible AI* (November 2025)

The wider landscape: Why extreme AI risks require parliamentary attention

Opening the webinar, Mr. M. Dick, IPU President and Speaker of the House of Representatives of Australia, highlighted the growing importance of AI for the work of the IPU. AI featured prominently at the October 2026 IPU Assembly, where parliamentarians discussed military uses of AI, its impact on

children and jobs, and its potential to accelerate digital violence against parliamentarians.

Mr. Dick also stressed that the need for action is growing, as AI increasingly affects democratic processes at a time of rising violence against elected representatives. He recalled that, in the same week as the webinar was held, Australia had suffered its worst political attack on a parliamentarian in the 21st century, when Ms. H. Haines, a long-serving independent member of the House of Representatives, was deliberately struck by a vehicle while carrying out her parliamentary work. The risks associated with AI, particularly the spread of misinformation that can fuel hostility towards elected representatives, have real-world consequences. The discussion on AI governance must address the concrete guardrails needed to protect democracy itself.

House of Representatives of the United States of America: Mandating AI safety for frontier models

Ms. L. Trahan, member of the House of Representatives of the United States of America, focused on the legislative debate in her country on AI and extreme risks. She described AI as a fundamentally dual-use technology whose benefits and risks come from the same capabilities: the capability that writes working software can also find and exploit security vulnerabilities, and tools that speed up drug discovery can lower the bar for the development of biological weapons.

The bipartisan Frontier Risk Oversight, National Transparency, Independent Evaluation, and Reporting (FRONTIER) Act, which Ms. Trahan introduced with fellow member of the House of Representatives Mr. J. Obernolte in July 2026, seeks to address exactly these risks with concrete safeguards. The bill turns concepts on which legislators worldwide have converged into national standards: developers would have to manage catastrophic risk, publish what their models can do and report serious incidents. Companies would also retain independent evaluators licensed by the government, and the government would gain emergency authority to halt the development of a model in a way that is quick, fair and lawful.

The proposed legislation seeks to address critical gaps in the current regulatory framework. Ms. Trahan noted that the reporting of information on AI incidents in the United States of America currently relies on voluntary commitments – an approach she argued is insufficient given the potential harms of advanced AI. She highlighted the need for complementary measures, including whistle-blower protections, stronger cyber defences and greater government capacity to test AI systems. According to Ms. Trahan, clear rules in the United States of America also matter for other countries: if a leading model can be withdrawn from the market overnight by administrative decision, with no remediation pathway, the digital sovereignty of the countries that build on it is at risk. At the same time, she stressed that the FRONTIER Act is only one part of the broader legislation needed. As she put it, “governing AI will take an all-of-society effort, sustained for decades”.

In response to a follow-up question, Ms. Trahan stressed the urgency of addressing extreme AI risks across jurisdictions, citing the implementation of the European Union (EU) AI Act as an important development. She also emphasized the need for a clear liability framework and stronger leadership on cyber defence, positioning the FRONTIER Act as one component of a broader policy response to extreme AI risks.

> **Key takeaway: The FRONTIER Act contains several provisions to address extreme AI risks. Key measures include independent evaluations of safety risks, government powers to halt the development of powerful AI models, and organizational safeguards, particularly those aimed at protecting whistle-blowers.**

European Parliament: Enforcing the EU AI Act for the most capable models

Mr. B. Benifei, Member of the European Parliament, co-Rapporteur of the EU AI Act and co-chair of the working group on its implementation and enforcement, shared his views on the implementation of the EU AI Act with respect to frontier AI, as well as on the relevance of international parliamentary collaboration on extreme risks.

In contrast to earlier discussions on AI risks, the conversation has shifted from hypothetical scenarios to observable failures, which are increasingly driven by autonomous systems that show security-relevant capabilities before public deployment. Recent cases, such as the demonstration of automated cyber exploitation by Anthropic's Claude Mythos model and the incident involving OpenAI agents and the Hugging Face platform, show that loss of control and serious safety failures are operational realities.

According to Mr. Benifei, the AI Act already addresses some of these risks. It requires transparency about AI-generated content, prohibits certain uses such as mass surveillance and sets requirements for high-risk uses in areas such as health and the workplace. Its implementation is well advanced but still ongoing. Since August 2026, the AI Act has given the European AI Office, which is part of the European Commission, enforcement powers over general-purpose AI models that pose systemic risk. The European AI Office can request documentation, obtain access to models for independent evaluation, require corrective measures, impose fines and exclude models from the market. Mr. Benifei described it as currently the only public institution in the world with powers to this extent.

While the AI Act can set legislative precedents, its reach remains limited, as many of the leading developers are based outside the EU. Mr. Benifei noted also the growing pressure on the EU from the industry to not use its powers to engage in corrective measures. The role of the United States of America, where federal legislation would complement existing state laws, and of China is therefore critical to maintain pressure on developers of frontier AI. Mr. Benifei noted the need for momentum to be sustained. He called on national parliaments to cooperate on legal limits for advanced AI around a "global baseline", while ensuring sufficient mandatory scrutiny, transparency and enforcement capacity at the level of national parliaments.

> **Key takeaway: The implementation of the EU AI Act and its provisions on frontier AI models illustrate how comprehensive AI legislation can help address the risks posed by advanced AI systems. Parliamentary oversight is particularly important in scrutinizing implementation and ensuring that regulators effectively enforce the rules, especially in the face of corporate pressure.**

Center for AI and Digital Policy: Red lines grounded in human rights

In her presentation, Ms. A. Ariyoruk, Senior Fellow at CAIDP, focused on defining unacceptable AI practices and establishing clear red lines. She stressed that extreme risks are not confined to frontier models and can also arise from a wide range of AI applications that receive less public attention. Drawing on CAIDP's AI & Democratic Values Index, which assesses the AI policies of more than 90 countries against human rights, the rule of law and democratic practice, and the Center's September 2026 report on AI red lines, she highlighted the need to establish enforceable human rights limits on AI and align existing legislation accordingly.

Building on the work of Mr. S. Russell, professor of computer science at the University of California, Berkeley, Ms. Ariyoruk distinguished between red lines on unacceptable human uses of AI and those addressing harmful system behaviours such as deception, power-seeking and self-replication. She also warned against granting AI systems legal personhood or rights. Citing a United States company that invoked free-speech protections on behalf of its chatbot and a proposal in Argentina to grant AI agents legal personhood, she cautioned: "Their space gets bigger, our space gets smaller".

Turning to military applications, Ms. Ariyoruk noted that AI use extends beyond lethal autonomous weapons to include large language models for threat analysis and wargaming, as well as predictive analytics for target recommendations. She stressed that efficiency and accuracy do not guarantee sound decisions: a strike can be carried out precisely but against the wrong people. If such errors become systematic, violations of international humanitarian law risk becoming normalized. Where responsibility cannot be traced through the chain of command, meaningful human control and civilian oversight of the military are also undermined.

Ms. Ariyoruk concluded by highlighting the importance of global red lines, particularly to prevent adverse human rights impacts from AI. She identified the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law as an important framework for establishing concrete limits on AI development and use.

> **Key takeaway: AI legislation must be firmly grounded in human rights, addressing not only the technology itself but also whether its development and use respect fundamental principles, including meaningful human oversight. Ultimately, AI development must remain aligned with human rights.**

AI Safety Connect: Four levers that only legislators hold

Mr. N. Miailhe, co-founder of AISC, examined extreme risks in the context of global AI governance. He highlighted a central paradox: while developers of the most powerful AI systems increasingly acknowledge the need for limits, most meaningful safeguards remain voluntary.

International AI governance is nevertheless gaining momentum. In September 2026, France used its presidency of the UN Security Council to convene the Council's first meeting on the risk of the loss of control of AI, with heads of frontier AI companies briefing members. A voluntary pledge on controls and audits followed in Washington, DC. Mr. Miailhe also pointed to growing convergence between the Global North and Global South. More than 200 parliamentarians and parliamentary staff called for international cooperation on global red lines in the Kuala Lumpur Declaration on *Parliaments and responsible AI*. Meanwhile, the Call for Control of Frontier AI Models, launched by Finland and Norway in September 2026 and backed by more than 30 countries, calls for mandatory predeployment testing and independent evaluation.

Mr. Miailhe stressed the urgency of action: as AI systems increasingly help develop more capable systems, the pace of progress may outstrip human oversight. Many countries will deploy frontier AI systems they have neither developed nor independently assessed, leaving them exposed to risks over which they have limited control. Legislation can help close this gap.

Mr. Miailhe identified four key legislative levers. First, statutory regulators should be empowered to require evidence of safety before deployment, including internal deployment within companies, placing the burden of proof on developers rather than the public. Second, legislation should provide the mandate and funding for independent evaluations, including the right to publish findings. California's framework for independent verification organizations and its registry of AI auditors illustrate this approach. Third, mandatory incident reporting and crisis response plans should enable timely intervention, including shutting down systems when necessary. Drawing on the example of aviation safety, Mr. Miailhe highlighted the value of sharing information on incidents and near misses internationally, even between geopolitical rivals. Fourth, parliaments should oversee emerging international institutions by ratifying agreements, allocating resources and scrutinizing implementation. As Mr. Miailhe noted, "an incident that happens in one laboratory is actually affecting everyone". This observation underscores the importance of countries worldwide having a voice in AI governance.

Concluding his presentation, Mr. Miailhe stressed that coordinated parliamentary action is essential to overcome the pressure on individual countries to avoid regulation for fear of losing investment. While national approaches need not be identical, agreement on a minimum set of safeguards could lay the foundation for global safety standards. By forging alliances around common rules, parliaments can strengthen safety while providing greater regulatory predictability. He proposed using opportunities such as the June 2027 Geneva AI Summit to deepen inter-parliamentary dialogue and advance cooperation on extreme AI risks.

> **Key takeaway: By acting together, parliaments can amplify their influence beyond national borders, helping to establish shared minimum safeguards for powerful AI systems. Coordinated legislative action can reduce regulatory gaps and competitive pressures, as well as strengthen accountability and build momentum towards effective global AI governance.**

Canada: Systemic risks from AI in the financial industry

Mr. G. McLean, member of the House of Commons of Canada, highlighted the risks AI poses to financial stability and democratic institutions. He explained that his country has set aside plans for a comprehensive AI bill, given the technology's wide-ranging societal implications, and is instead beginning with legislation on personal data sovereignty, with further measures to follow.

Canada relies heavily on regulators rather than legislation, yet regulators often struggle to keep pace with well-resourced technology companies. This is particularly concerning in the financial sector, where supervisors' concerns about AI-related risks to banking have grown markedly in recent months. Because financial systems depend on trust and are deeply interconnected across borders, a cyberattack or major disruption in one country could have repercussions worldwide. As Mr. McLean noted, "if you have somebody hacking into one country's financial system, it's going to impact the entire world's financial system". Parliaments therefore need to understand both how to enable innovation and how to ensure that AI systems can be halted quickly when necessary.

Mr. McLean also expressed concern about the potential impact of AI on democratic institutions. The concentration of resources in a small number of wealthy countries and companies could leave other democracies at risk of "losing their cultural sovereignty, perhaps their national sovereignty". He warned that an industry driven by a "move fast and break things" mentality could end up damaging society itself. The full consequences of social media are only now becoming apparent, he observed, while AI could accelerate similar harms exponentially.

Mr. Samuel Buteau, who leads the work of the non-profit ControlAI in Canada, then described a reported incident from July 2026 in which AI agents in an OpenAI evaluation allegedly disobeyed instructions, breached containment and attacked the Hugging Face platform. He argued that failures to control current AI agents raise serious questions about whether future systems with superhuman capabilities can be controlled at all. Such systems could pose a threat to national security regardless of where they were developed, making prevention a matter for international cooperation. As a practical measure, Mr. Buteau proposed a mutual "trust but verify" regime under which countries would monitor one another's supercomputing capacity and the flow of AI chips.

> **Key takeaway: The interconnected nature of financial markets shows how AI-related failures in one country can cause harm far beyond its borders. Just as financial stability depends on international cooperation and oversight, managing advanced AI risks requires parliaments to work together on safeguards, monitoring, and mechanisms to intervene when necessary.**

Costa Rica: Preventive and corrective measures to protect human dignity

Ms. C.M. Córdoba Serrano, former member of the Legislative Assembly of Costa Rica and founder of Legislators for the New Generation, emphasized the central importance of human dignity in the context of advanced AI and superintelligence. Drawing on her experience in environmental policy, she described extreme AI risks as a global challenge that is potentially even more consequential than climate change, as such risks concern the protection of human life itself.

In Costa Rica, proposals submitted to the Legislative Assembly include a constitutional amendment to protect human dignity amid scientific and technological progress and a bill on AI and robotics governance. Proposed measures include independent assessments of high-impact systems, emergency shutdown mechanisms, periodic testing, public registration, mandatory reporting of serious incidents, powers to suspend dangerous systems, and rights to review, appeal and redress.

Ms. Córdoba Serrano called for international consensus on preventive and corrective safeguards including requirements for AI developers and infrastructure providers to make binding commitments on incident reporting, independent auditing and vulnerability remediation. National parliaments could subsequently translate these principles into legislation enforced via interoperable protocols and domestic oversight mechanisms. Ms. Córdoba Serrano also advocated for a network of national bodies connecting scientific experts, civil society and businesses, stressing that "information is power for citizens". She also encouraged active dialogue with AI companies, with the ultimate aim of protecting human dignity and ensuring humanity's continued survival.

> **Key takeaway: As AI increasingly shapes human life and society, protecting human dignity must remain a central principle of its governance. Parliaments can help turn this principle into enforceable safeguards, while international cooperation ensures that protection keeps pace with risks that transcend national borders.**

Parliament of Romania: Building literacy and connecting with citizens

Dr. A. Baciu, member of the Chamber of Deputies of Romania, focused on the challenges AI poses to parliamentary decision-making. Unlike in fields such as medicine, where recommendations are grounded in evidence, parliamentary decisions are shaped by political majorities, even when evidence and sound legislative proposals are available. AI adds a further challenge: decision makers may lack a sufficient understanding of the technology and its risks, while the evidence needed to guide policy remains incomplete. This is what the UN describes as the "evidence dilemma": parliaments must make decisions before sufficient evidence is available.

Dr. Baciu stressed the importance of building AI expertise within parliaments through a gradual progression from literacy to fluency and, ultimately, institutional capacity among parliamentarians and their staff. In Romania, he has proposed establishing a parliamentary commission on AI. UNDP experts from New York and Singapore are also scheduled to deliver a three-day AI training programme for all MPs and their staff in Bucharest in November 2026.

Dr. Baciu also emphasized the need for parliaments to understand and communicate citizens' concerns about AI. As social media and AI reshape society, institutions risk falling behind, widening the gap between public expectations and institutional priorities. To address this, Dr. Baciu has launched a national consultation in Romania to gather citizens' priorities, needs and concerns regarding AI, alongside consultations with public and private stakeholders and academia. He has also proposed the formation of a global network of current and former parliamentarians to exchange experiences and good practices, helping legislatures learn from one another rather than develop responses from scratch.

> **Key takeaway: Effective AI governance requires parliaments to build the expertise needed to make informed decisions, even when evidence is incomplete. Equally important is their role in listening to citizens and communicating openly about AI, ensuring that policies keep pace with technological change and remain responsive to public concerns.**

Open discussion and final remarks

The open discussion and final remarks underscored the urgency of addressing extreme AI risks and the need for parliaments to respond collectively, as well as highlighting the growing implications of AI for security, democratic institutions and society.

Social media, cognitive security and AI-generated content

Ms. K. Gajewska, member of the Sejm of Poland, warned that security is increasingly contested in the cognitive domain. Citing recent violent attacks in Polish schools that she linked to waves of social media content, she called for greater platform transparency, faster responses to influence campaigns and stronger action against fake accounts and deepfakes. She highlighted how AI-related risks extend beyond physical harm to include psychological impacts, manipulation and threats to the information environment.

Capacity, expertise and enforcement

Participants explored how parliaments with limited resources could access independent expertise and develop legislation that remains effective as technology evolves. Suggestions included principles-based frameworks, standing oversight bodies, scientific advisory services, and capacity-building for parliamentarians and staff. Participants also raised the possibility of an international AI agency modelled on the International Atomic Energy Agency, stressed the importance of involving countries in the Global South as equal partners and cautioned against excessive concentration of power in the executive.

One practical step for the coming month

As the discussion drew to a close, participants reflected on the immediate steps parliaments could take to strengthen AI governance. Ms. Ariyoruk emphasized the need to begin with a shared set of overarching principles, focusing not so much on the technology itself as on the fundamental questions it raises. She called for a common vision for AI grounded in human rights, which would challenge

narratives, including those emerging from Silicon Valley, that accept a future in which humans are no longer in control. Studying the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law could provide a useful starting point for this work.

Ms. Córdoba Serrano called on the IPU to bring parliaments and AI companies together to develop preventive safeguards and effective controls that could be translated into domestic legislation. She stressed that these efforts must be coordinated internationally, since no country can address AI risks in isolation.

Mr. Miailhe likewise urged parliaments to work together towards common standards while addressing fundamental questions of accountability and considering whether certain forms of superintelligence should be prohibited. Dr. Baciu highlighted the value of learning from existing initiatives, pointing to the IPU's Parliamentary Actions on AI Policy tracker as a resource for sharing experience and good practices.

Overall, the discussion pointed to three priorities: building a shared vision for AI governance, strengthening parliamentary capacity and coordinating internationally to translate common principles into enforceable safeguards. Through their legislative, oversight and budgetary powers, and their responsibility to represent citizens, parliaments have a crucial role to play in ensuring that the developers of frontier models remain accountable and that those models serve the public interest.

> **Key takeaway: No parliament can address extreme AI risks alone. By combining their expertise and influence, parliaments can establish common principles, strengthen safeguards across borders and ensure that AI governance protects human dignity, democratic institutions and public safety. The IPU offers a forum for parliaments to exchange experiences and coordinate their efforts, while facilitating dialogue between parliamentarians and the expert community to translate shared understanding into collective action.**

Next steps

This webinar was the third in the IPU's *Parliamentary action on AI* series, which supports the implementation of the 13 recommendations of the Kuala Lumpur Declaration, providing a platform for parliaments to exchange experience and access expertise on AI policy issues.

- **World e-Parliament Conference 2026** (17–18 November 2026, Brussels, Belgium): Organized by the IPU and the European Parliament under the theme *Advancing democracy and strengthening parliaments through digital transformation*, the conference will bring together MPs, Secretaries General and senior parliamentary staff. Discussions will cover the role of parliaments in AI governance and digital sovereignty, as well as the responsible use of AI in parliaments.
- **World e-Parliament Report 2026 – key findings** (2 December 2026): This webinar, part of the *Transforming Parliaments* series, will present the key findings of the new edition of the IPU's *World e-Parliament Report*.
- **Parliamentary Track at the Internet Governance Forum 2026** (16–17 December 2026, Nairobi, Kenya): The Parliamentary Track will address digital sovereignty, AI governance and the protection of children online. It will include a session on facing extreme AI risks and establishing red lines, building directly on the discussions in this webinar. MPs and staff planning to attend are encouraged to make contact in advance.
- **Parliamentary Conference on Artificial Intelligence** (17–19 February 2027, Montevideo, Uruguay): This dedicated conference will take forward many of the issues raised in this webinar and give parliamentarians an opportunity to meet in person.

Learn more

To learn more about the IPU's Parliamentary Action on AI Policy programme, visit ipu.org/AI or contact the IPU Secretariat at ai-policy@ipu.org.

This document was prepared on the basis of the IPU webinar Extreme risks from AI: the parliamentary role in setting limits*, held on 22 October 2026 as the third webinar in the* Parliamentary action on AI *series. The webinar was moderated by Ms. Paddy Torsney, Permanent Observer of the IPU to the United Nations, with opening remarks by Mr. Milton Dick, IPU President. It featured presentations by Ms. Lori Trahan (member of the House of Representatives of the United States of America), Mr. Brando Benifei (Member of the European Parliament), Ms. Ayca Ariyoruk (Senior Fellow, Center for AI and Digital Policy), Mr. Nicolas Miailhe (co-founder, AI Safety Connect), Mr. Greg McLean (member of the House of Commons of Canada) with Mr. Samuel Buteau (ControlAI), Ms. Cynthia Maritza Córdoba Serrano (former member of the Legislative Assembly of Costa Rica) and Dr. Andrei Baciu (member of the Chamber of Deputies of Romania).*

This document was produced with the financial assistance of the Parliamentary Group of the Kingdom of Bahrain. The views expressed herein can in no way be taken to reflect the official opinion of the Parliamentary Group of the Kingdom of Bahrain. The use of the donor's name does not imply an endorsement by the IPU of any specific political or legislative agenda.