

Moonshots in Reliable, Safe and Secure AI: A Call for European Leadership


1. We, the undersigned policymakers, industry leaders, researchers, and civil society representatives gathered at the European Parliament, share a common conviction: the European Union has a rare, time-limited strategic opportunity to lead the world in **Reliable, Safe and Secure Artificial Intelligence (RSS AI)**. We call on the EU to seize this opportunity.

A CRITICAL GAP AT THE HEART OF THE AI RACE

2. The global AI race is accelerating. In 2025 alone, the world's major technology powers committed hundreds of billions of dollars to scaling AI infrastructure. Yet, this race is being run without adequate attention to a foundational problem: **today's most powerful AI models cannot offer formal guarantees to do what we need them to do where it truly matters**. We cannot ensure that AI systems will consistently act in accordance with user intent (**reliability**), operate within defined risk boundaries (**safety**), or withstand adversarial exploitation (**security**).
3. These limitations are barriers that prevent AI from being reliably deployed in critical infrastructure sectors: energy grids, rail networks, aerospace, healthcare, pharmaceuticals, or defence. Europe's champions (companies such as Airbus, Siemens, Alstom, Philips, or Thales) are heavily concentrated in these safety-critical sectors compared to their US and Chinese counterparts. Therefore, **for Europe, unreliable AI is not a technical inconvenience or theoretical concern: it is today's barrier to adoption**.
4. **Leading on RSS AI research and development would unlock whole categories of AI applications that are currently out of reach, boost competitiveness**, by delivering productivity gains in leading European sectors, and **create strategic benefits for Europe** in the global AI context.

A STRATEGIC OPPORTUNITY EUROPE MUST SEIZE NOW

5. RSS AI research and development is severely under-financed globally. Frontier AI companies face competitive dynamics that crowd the fundamental research required for RSS. In addition, whilst access to compute is a key component, the decisive inputs to lead on verifiably reliable AI are **world-class research talent and scientific ambition**, which makes it cheaper than competing on the compute intensive current scaling paradigm. Overall, the bar for Europe to become **the leading global investor and innovator in RSS AI** is therefore genuinely achievable, **but only if Europe acts with urgency**.
6. We identify various benefits to seizing this opportunity:
 - a. **Economic competitiveness**: RSS guarantees would unlock AI use cases for European industrial champions, allowing them to effectively seize AI-induced productivity gains.
 - b. **Strategic leverage**: leadership in the "RSS layer" of AI would confer strategic leverage to Europe, concentrating indispensable expertise to advise deployment around the world as reliable models become essential.

- 
- c. **Talent attraction and retention:** Europe faces a serious brain drain of AI expertise. Leading on RSS AI could draw researchers motivated to unlock some of the most useful AI applications.
 - d. **Global norm-setting:** a Europe that leads on RSS AI would export both requirements and solutions, reinforcing the Brussels effect and giving European AI regulation genuine technical credibility.

WHAT EUROPE NEEDS: A BOLD AND COORDINATED MOVE

- 7. We recognise and welcome the steps already taken through the AI Continent Action Plan, the InvestAI initiative, the Frontier AI Initiative, and the growing number of Member States' national AI programmes, as important foundations.
- 8. But critically, they remain fragmented. Investment in RSS AI is currently dispersed across disconnected national programmes, underfunded academic labs, and isolated civil society efforts, without the shared agenda, coordination, or scale needed to deliver breakthroughs. This fragmentation is Europe's greatest self-imposed obstacle: it dissipates talent, duplicates effort, and prevents the critical mass that serious moonshot research demands.
- 9. The opportunity before Europe demands a **decisive, pan-European intervention**, one that brings together policymakers, researchers, industry and civil society around a shared mission and a common research agenda.
- 10. We therefore call on the European Union and its Member States to:
 - a. **Establish a dedicated European ARPA-style research institution** for moonshot research in RSS AI, built for speed, scientific risk-taking, and mission focus, not for ad-hoc grant-making. Its mandate should be to pursue formal reliability guarantees, safe-by-design model architectures, and novel mechanisms for hardware and model security verification.
 - b. **Integrate RSS AI priorities into existing European initiatives**, including the Frontier AI Initiative, the AI Gigafactories, and the European Competitiveness Fund, ensuring that Europe's AI ambition is anchored in reliability as well as capability.
 - c. **Commit dedicated, multi-year funding** at the scale this ambition requires. Based on existing estimates, a serious foundational research programme would require on the order of **hundreds of millions of euros annually as a start**, several orders of magnitude less than competing on compute infrastructure, but concentrated, stable, and coordinated at the European level to achieve scale.
- 11. Europe can be the home of AI the world can rely on. We commit to working together, across sectors, disciplines, and national boundaries, to make it a reality.

Signatories

- MEP Brando Benifei, S&D
- Yoshua Bengio, Mila
- Jonathan Barry, Mila
- Axel Brando Guillaumes, Independent researcher
- Rachel Coldicutt, Careful Industries
- Lucy Czachowski, Bitkom
- Younesse Kaddar, University of Oxford & CAIA
- MEP Sergey Lagodinsky, Greens/EFA
- Anna Lukina, TU Delft
- MEP Michael McNamara, Renew Europe
- Nicolas Mialhe, AI Safety Connect
- Amin Oueslati, The Future Society
- Vanina Paolin-Gagin, Member of the French Senate
- Henry Papadatos, SaferAI
- Sam Ramadori, LawZero
- Max Reddel, Arq Foundation
- Charbel-Raphaël Ségerie, Centre pour la Sécurité de l'IA
- Patrick Stadler, Pour Demain
- Samuel Staton, University of Oxford & CAIA
- Max Tegmark, MIT and Future of Life Institute

This declaration reflects the conclusions of a roundtable held at the European Parliament on 23 March 2026, titled *Moonshots in Reliable, Safe & Secure AI: A Path to European Leverage and Strategic AI Adoption*. The event was hosted by MEPs Michael McNamara and Sergey Lagodinsky and co-organized by SaferAI and Pour Demain, bringing together around 40 policymakers, industry leaders, and researchers. A summary of the discussion is [available on the SaferAI blog](#). The declaration builds on the arguments developed in SaferAI's position paper, [The Case for European Investment in High-Risk, High-Reward AI Reliability Research](#).