

IASEAI Issues Call to Action for Lawmakers, Academics, and the Public Ahead of AI Summit in Paris

The International Association for Safe and Ethical Artificial Intelligence (IASEAI) concluded its inaugural conference today by calling for ten critical action items.

Feb. 7, PARIS: The [IASEAI](#) '25 conference has been the largest ever global gathering focusing on cutting-edge developments in AI safety and ethics. Held at the OECD headquarters just prior to the Paris AI Action Summit, the conference brought together leading representatives from academia, civil society, industry, media, and government.

Professor Stuart Russell, IASEAI President and Distinguished Professor of Computer Science at the University of California, Berkeley said:

"The development of highly capable AI is likely to be the biggest event in human history. The world must act decisively to ensure it is not the last event in human history. This conference, and the cooperative spirit of the AI Summit series, give me hope; but we must turn hope into action, soon, if there is to be a future we would want our children to live in."

The following actions, derived from the research presented and the deliberations held during the conference, will help to further IASEAI's mission to "ensure that AI systems are guaranteed to operate safely and ethically."

IASEAI calls for:

1. **Recognition of the significance of new developments in AI.** Policymakers must act with an urgency that matches the transformational potential of AI, the rapidity of change, and the increasing risks to humanity as AI capabilities begin to exceed our own.
2. **Preventing AI-driven institutional and social disruption.** The power of AI threatens to disrupt employment and social structures, worsen inequality, and severely compromise the information ecosystem. Policymakers must take proactive steps to protect institutions, individuals, and ways of life while harnessing AI to strengthen rather than weaken societies. To the extent that these problems take similar forms across nations, policymakers should collaborate to seek collective solutions.
3. **Addressing the race to AGI.** The financial and strategic incentives to achieve artificial general intelligence (AGI) or "superintelligence" lead companies and countries to undercut safety standards in an attempt to gain technological and political control. Policymakers should take coordinated actions in international fora such as the UN and the OECD to ensure that innovation and competition proceed within an agreed framework of rigorous safety standards.

4. **Coalescing the efforts of research communities around the goal of safe and ethical AI.** The increasing threats to human flourishing posed by AI systems require researchers from the AI, ethics, social science, and policy communities to collaborate and pool their efforts. The various perspectives offered by these communities are not in tension; on the contrary, they all contribute in important ways to ensuring that AI systems do not harm human society.
5. **Adoption of mandatory safety and ethical requirements.** While commendable, voluntary commitments by companies must be made more specific and legally binding. Such binding commitments might include mandatory registration of advanced AI systems, including automatic self-registration for open-source copies; installation of remotely operable off-switches; mandatory reporting of incidents; professional ethics training for engineers; and standards for training, design, development, testing, auditing, and transparency of advanced AI systems. Developers should provide a scientifically convincing safety case that their systems will not cross so-called *behavioral red lines*, which demarcate unacceptable behaviors.
6. **Advancement of global cooperation.** AI developments and the accompanying risks have cross-border impacts, necessitating global cooperation on AI safety research and regulation, considering the perspectives of all nations. Moreover, the benefits of AI must be equitably distributed. In this vein, we welcome the establishment of the UN AI Advisory Body and the international network of AI Safety Institutes.
7. **An increase in publicly funded research.** The scale of the challenge requires significantly more publicly funded AI safety and ethics research. Research cannot continue to be dominated by companies with significant conflicts of interest. The potentially transformative benefits of AI can be realized only if advances in capabilities are accompanied by methods to ensure that AI systems are safe by design and aligned with human interests.
8. **Support of the AI Foundation.** The forthcoming AI Foundation, as anticipated by the Paris Summit hosts, seems to be a significant step forward. The Foundation must support the design and development of AI systems to address human needs and enhance local capabilities, and governments must, in turn, support the Foundation.
9. **Support of the Council of Europe Framework Convention on Artificial Intelligence.** The Council's AI Treaty "aims to ensure that activities within the lifecycle of artificial intelligence systems are fully consistent with human rights, democracy and the rule of law." While the Council administers the treaty, it is open for signature by non-member states. 37 countries have signed, including the United States and United Kingdom, and more are expected to sign at the Paris Summit.
10. **Fostering informed dialogue.** AI researchers and the media need to contribute their complementary forms of expertise in informing the public and policymakers, objectively and accurately, about developments in AI and their significance.

About IASEAI: The International Association for Safe and Ethical AI (IASEAI) is a non-profit organization founded to address the risks and opportunities associated with rapid advances in AI. Their mission is to ensure that AI systems are guaranteed to operate safely and ethically; and to shape policy, promote research, and build understanding and community around this goal.