

Western and Chinese scientists: AI safety a “global public good”, global cooperation urgently needed

VENICE, ITALY, September 16, 2024 /EINPresswire.com/ -- Leading global artificial intelligence (AI) scientists gathered in Venice in September where they issued a call urging governments and researchers to collaborate to address AI risks.

Computer scientists including Turing Award winners Yoshua Bengio and Andrew Yao, as well as UC Berkeley professor Stuart Russell, OBE and Zhang Ya-Qin, Chair Professor at Tsinghua University, convened for the third in a series of International Dialogues on AI Safety (IDAIS), hosted by the Safe AI Forum (SAIF) in collaboration with the Berggruen Institute.

The event took place over three days at the Casa dei Tre Oci in Venice and focused on safety efforts around so-called artificial general intelligence. The first day involved a series of discussions centered around the nature of AI risks and the variety of strategies required to counter them. Session topics included early warning thresholds, AI Safety Institutes, verification and international governance mechanisms.

These discussions became the basis of a consensus statement signed by the scientists, centered around the idea that AI safety is a “global public good”, suggesting that states carve out AI safety as a cooperative area of academic and technical activity. The statement calls for three areas of policy and research.

First, they advocate for “Emergency Preparedness Agreements and Institutions”, a set of global authorities and agreements which could coordinate on AI risk. Then, they suggest developing “Safety Assurance Frameworks”, a more comprehensive set of safety guarantees for advanced AI systems. Finally, they advocate for more AI Safety funding and research into verification systems to ensure that safety claims made by developers or states are trustworthy. The [full statement can be read here](#).

On the second day, scientists were joined by a group of policymakers, former President of Ireland Mary Robinson and other experts. The scientists emphasized the urgency of employing these proposals given the rapid pace of AI development. The statement was presented directly to the policymakers, and the group strategized about how the international community may work together to accomplish these goals.

International Dialogues on AI Safety

Safe AI Forum

idais@far.ai

Visit us on social media:

[X](#)

[LinkedIn](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/743901916>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2024 Newsmatics Inc. All Right Reserved.