

**FUTURE
INVESTMENT
INITIATIVE**
Powered by the FII Institute



AI RESOLUTION



King Abdulaziz International Conference Center
Riyadh, Saudi Arabia

AI RESOLUTION

Today's rapidly advancing AI models are transforming the global economy and empowering small teams to advance science and technology and uplift humanity.

Driving these advances are the rapid advancement of computation power, the exponential growth in the size of labelled data, dramatic reductions in training costs, a surge in capital going into AI development, and democratized access advanced AI models.

The result is a lower barrier to entry that will accelerate adoption, increasing both the potential of AI for good and its use by malevolent players to cause disruptions on a global scale.

Many of the near-term global risks are known but few have been fully thought-through. Over the next five years, these risks include the use of AI for misinformation, cyber-attacks, political and financial disruption, as well as the creation of bioweapons. Few if any of these have an adequate response plan agreed upon by political and corporate leaders,

Medium-term risks also abound, from the impact of AI on societal structures and jobs to the risk of Artificial General Intelligence (AGI), which many leading experts expect within the next 5-15 years. If misaligned with the needs of humanity, such AGI (and the Artificial Super Intelligence that is expected to come after that) could create existential risks for mankind.

The above global risks are considered by most to be likely, and the key question is thus the degree of the impact and our level of preparedness. World leaders must be prepared to act with the appropriate governance and policies.

AI RESOLUTION

RECOMMENDATIONS TO GOVERNMENTS:

1 Promote AI Alignment

The speed at which AI is developing requires a more dynamic approach to policy. Given AI's future capabilities and associated risks are hard to predict, governments should:

- **Require all AI companies to document and disclose how they are committed to and are implementing AI alignment.** Public AI funding and procurement of AI-related services should be reserved for companies that are committed to AI Alignment and have a demonstrated implementation plan to maximize such alignment.
- Work collaboratively with industry and a multi-lateral national leadership to **develop a set of prerequisites for the release of AI models and support the development of early warning systems** that can track, evaluate, and report on the risks of those models once deployed.
- **Require AI developers to run comprehensive model assessments prior to release and make the assessments results publicly available.** These assessments should provide full transparency on their AI Models, and importantly, how their AI models are being trained. Governments should also **work with private industry to develop an accepted set of training datasets considered safe and pro-alignment.**
- **Increase funding for AI alignment research at the highest level** and the development of new AI paradigms that are explicitly safe by design (e.g. provable AI).

2 Establish a Global and Inclusive AI Governance Regime

Global risks associated with AI can only be addressed through a global approach. Right now, however, most existing international coordination mechanisms lack appropriate representation of emerging markets. A new international regime is thus urgently needed, and it should be inclusive of both developing and developed markets.

3 Prepare For The Transition Today

- **Actively invest in upskilling /reskilling programs and social safety nets** to ensure a just transition for workers displaced by AI.

AI RESOLUTION

RECOMMENDATIONS FOR THE PRIVATE SECTOR & INVESTMENT SECTOR:

It is important to recognize that the private sector is under tremendous competitive pressure to maximize the power of its AI models and release them for consumption by their users as rapidly as possible. Companies and investors have a responsibility in ensuring this competition drives a race towards safer AI, not an AI race to the bottom.

1 Foster Transparency & Accountability

- Commit to AI alignment and create alignment plans detailing the strategies and processes that will be implemented to deliver such alignment.
- Ensure new models come with publicly available detection mechanisms that enable people to know whether a particular content is AI generated;
- Report failures or risks when identified and share best practices on how to address them

2 Actively Reduce AI Risks

- Collaborate with governments to develop adequate global response capacity to the known global risks caused by existing AI models;
- Adopt a safety-level based approach to development of future AI, defining development procedures commensurate with the safety-levels and engaging in conditional slowdown when necessary;
- Ensure the security of State-of-the-Art AI facilities and systems

INVESTMENT

- The Investment Community (VC, PE, Family Offices, R&D funding) and other sources of funding should commit to only providing AI financing to AI-related companies that have proven AI-Alignment policies in place.

AI RESOLUTION

OUR FII COMMITMENT:

To support the public and private sectors towards the implementation of these recommendations, the FII Institute endeavors to become a globally leading platform to:

1 **THINK on AI**

By working on solutions that can be adopted by all and building a common understanding of how this technology can help to tackle some of our greatest problems.

2 **ACT on AI**

by investing to scale up global response capacity. Solutions like detection models or those that help with AI interpretability are particularly promising as they can help across multiple risks domains.

3 **XHANGE on AI**

By establishing the FII AI Coalition, a unique and neutral platform bringing together investors, corporations, government leaders and academics from developed and emerging markets to engage in a frank and candid dialogue on the critical issue of AI alignment.

4 **SUPPORT AI ALIGNMENT**

By committing to fund only those companies that have demonstrated a full commitment to AI Alignment - and a plan for how they intend to maximize such alignment.

Over the coming months, we will be bringing the findings and thinking of the FII AI Coalition to all international efforts focused on concrete responses.

We look forward to working with all interested parties - because this concerns us all and only by working together can we ensure AI positively impacts humanity.



**FUTURE
INVESTMENT
INITIATIVE**
Powered by the FII Institute



For more information, please visit our website: www.fii-institute.org

Follow us:   @FIKSA     FII Institute