Science & Technology · GS - III

Advanced artificial intelligence offers significant socio-economic benefits but may also create risks that extend beyond national boundaries. Discuss the need for a comprehensive AI safety governance framework.

EXAM YEAR	PAPER	SUBJECT
DAILY PRACTICE 28/11/2026	GS - III	Science & Technology

MARKS	WORD LIMIT
15	~250 words

Answer Hint

AI safety, catastrophic risks, dual-use technology, biosecurity, cybersecurity, accountability, human oversight, international cooperation

Model Answer

Artificial Intelligence is transforming healthcare, education, industry and governance. However, increasingly capable AI systems also create risks that may affect public safety, national security and global stability.

The possibility of misuse or loss of control over advanced systems has made AI safety an important governance priority.

Major areas of concern include:

1. **Dual-use applications:** AI tools developed for legitimate research may also be misused to facilitate cyberattacks or biological threats.
2. **Information integrity:** Generative AI can enable deepfakes, automated misinformation and manipulation of public discourse.
3. **Cybersecurity vulnerabilities:** AI may increase the scale and sophistication of cyber threats against critical infrastructure.

4. **Loss of human oversight:** Excessive dependence on automated systems in sensitive sectors can create accountability gaps.
5. **Systemic risks:** Highly interconnected AI applications may produce cascading failures across financial, administrative and technological systems.

A comprehensive governance framework should incorporate:

- Mandatory risk assessments and independent safety evaluations before deployment of high-risk systems.
- Red-teaming, continuous monitoring and incident-reporting mechanisms.
- Clear accountability for developers, deployers and institutions.
- Strong safeguards for sensitive biological and cyber capabilities.
- International cooperation on safety standards and information sharing.

India can strengthen this framework through its IndiaAI Mission, responsible innovation policies and coordination among regulatory and research institutions.

Ultimately, AI governance should neither obstruct beneficial innovation nor allow technological advancement to outpace safety. A risk-based, transparent and internationally coordinated approach is essential.

BECOME IAS
AI POWERED UPSC PREP