



POSITION STATEMENT

EFFECTIVE GOVERNANCE OF ARTIFICIAL INTELLIGENCE

*Adopted by the IEEE-USA
Board of Directors (17 November 2023)*

IEEE-USA believes that while AI systems¹ contribute to human and societal well-being, to realize the benefits, this technology must be ethically designed and adhere to principles that respect human rights, privacy, transparency, accountability, and equity while simultaneously encouraging innovation and effectiveness. By providing the right balance of investments, incentives, frameworks, and regulation, those charged with AI governance must guarantee that current and future systems maximize their social benefits, while mitigating potential harms.

Our recommendations for governing AI fall into three groups. **First**, prioritize the design, development, and deployment of competitive, safe, trustworthy, and ethical systems that are aligned with democratic values and with the protection of societal well-being. **Second**, ensure that the U.S. has an AI-capable workforce to realize the economic benefits of this technology, while adhering to meaningful democratic governance.² **Third**, generate and implement risk-mitigating frameworks that prevent harms such as controversies, critical failures, and even the loss of life without hampering innovation.

1. To encourage the design, development, and deployment of competitive, safe, trustworthy, and ethical AI systems, IEEE-USA recommends that the U.S. government:

- a) **Ensure awareness, access, research, and testing of AI systems.** Any AI system with a likelihood of generating individual or societal harm, regardless of its use context, must be properly vetted and those responsible should be held accountable.³ For this to happen, the government should clarify whether and how proprietary AI systems may be reverse engineered, modified, and evaluated under federal laws (and their state equivalents). More broadly, it should take steps to affirmatively promote awareness, access, research, and testing including:

¹ An AI system is a machine-based system designed to operate with varying levels of autonomy which can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. This includes computer systems and networks that can perform tasks which traditionally required human intelligence such as visual perception, speech recognition, learning, decision-making, control, and natural language processing. The term AI also includes automated decision systems.

² IEEE-USA, Artificial Intelligence: Accelerating Inclusive Innovation by Building Trust. IEEE-USA, 2020. Available: <https://ieeusa.org/assets/public-policy/positions/ai/AITrust0720.pdf>

³ *How Should We Regulate AI: Practical Strategies for Regulation and Risk Management from the IEEE1012 Standard for System, Software, and Hardware Verification and Validation*; IEEE-USA, August 2023. <https://ieeusa.org/product/how-should-we-regulate-ai/>

- Ensure accountability and transparency in government procurement and contracting for AI systems;
 - Introduce legislation that holds the designers, developers, and/or deployers of AI systems accountable for individual and group harms and prevent immunity to liability;
 - Adopt clear governance relating to the collection, usage, storage, and sharing of personal information by private and public stakeholders;
 - Assess optimal approaches for the deployment of open-source AI systems and propose holistic alternatives to govern this technology in a manner that balances the strengths of the cooperative nature of open source with the risks of providing access to harmful AI capabilities that are scalable.
 - Require organizations to conduct AI audit assessments to understand and document the impact of automated decision systems on internal and external stakeholders;
 - Develop mechanisms for identifying and accounting for the features of AI systems that could cause current testing, evaluation, certification, and investigation methods to misinform decision makers or the public about the risk of system deployment or the causes of system malfunction; and,
 - Specifically, in legal disputes, tribunals should permit disclosure under appropriate protective orders of intellectual property related to AI systems when necessary to obtain evidence in compliance with other judicial requirements, including constitutional requirements, discovery laws, or subpoenas.⁴
- b) **Increase investment in AI to remain competitive.** A bi-partisan priority must be the investment in AI research, development, and infrastructure to help expand innovation, while also maximizing societal benefits and mitigating any associated risks. Doing so will lay the groundwork for a strong domestic AI industry that continues to innovate by increasing the quality and quantity of large, medium, and small AI entities.
- c) **Develop and promote standards that adhere to the core principles of due process, openness, consensus, balance, and the right of appeal.** Standardization provides a number of benefits, including setting the foundation on which technology innovation is based, enabling the development of complex solutions at a better cost structure, fueling growth of global markets, expanding consumer choice, supporting interoperability, and helping protect the health and public safety of workers and the general public. To ensure these benefits are realized, the government should prioritize the development, promotion, and adherence to standards for ethical and trustworthy AI systems. Moreover, they should require that such standards be generated through a market-driven consensus-based development process that adheres to the core principles of due process, openness, consensus, balance, and the right of appeal.
- d) **Improve its understanding of the data provenance and outputs of AI systems throughout their supply chain and life cycle.** AI tools such as large language models (LLMs) have the potential to result in logical and mathematical inaccuracies in text, incorrect or impossible images, and unacceptable or hurtful biases. Understanding these unique threats requires thoughtful guardrails to be put in place to prevent risk across the model building lifecycle (pre and post deployment). This raises the need to examine how data is obtained and processed at the multiple phases of collection.

⁴ See, IEEE-USA, *Artificial Intelligence: Accelerating Inclusive Innovation by Building Trust*, supra note.

2) To realize the economic benefits of AI systems through an AI-knowledgeable and capable workforce, IEEE-USA recommends that the U.S. government:

- a) **Calibrate public trust, understanding, and discourse about AI systems.** Responsible development and governance of AI systems requires ensuring transparency in how this technology makes decisions, as well as public understanding of its capabilities and limitations. To achieve this, the U.S. should develop public outreach strategies that inform about AI, improve broad and inclusive participation in its design and regulation, and develop appropriate levels of trust in the technology. In particular, the public must understand and be aware of the following:
- When particular AI techniques are in use;
 - AI systems' competency and the extent to which they might produce disparate impacts;
 - Whether they are safe and secure, and how this is evaluated;
 - Their legality and legal accountability;
 - Their impacts on privacy;
 - Whether they might constrain the autonomy of users or other affected individuals; and,
 - Their potential ethical and societal impacts.
- b) **Support AI education and retraining opportunities to meet current and future workforce needs.** The extraordinary growth in AI has created demands for talent in industry, government, and academia. Internationally, fierce competition exists for AI expertise, leading to the threat of a loss of talent from the U.S. workforce. To maintain its technological competitiveness, the U.S. must encourage students to pursue degrees in AI-related disciplines (within and outside of the STEM pipeline), attract and retain international students by reducing uncertainty in immigration policy, and develop AI-related continuing education programs. Addressing workforce needs will help to maintain U.S. competitiveness internationally and ensure that the workforce possesses relevant skills in the future.

3) To generate sensible governance frameworks for AI, IEEE-USA recommends that the U.S. government:

- a) **Establish reliable, predictable, and robust governance tools for AI systems.** Methods and applications of AI will continue to push the boundaries of regulation and social norms for the foreseeable future. To address these issues, institutions should **harness a diverse toolkit of strategies to manage the technology.** For instance, when flexibility and experimentation are of value, soft law (i.e. programs that create substantive expectations, but are not directly enforceable by government) in the form of guidance, principles, and standards, among others, can serve as a basis to evaluate and identify effective solutions. However, flexibility and experimentation should not be interpreted as allowing a suspension of ethical principles. When enforcement is prized, regulation or hard law at the national or local level should be designed with the input of stakeholders in all sectors. Reliable, predictable, robust, and adaptable frameworks that are written in accordance with recognized U.S. and international legal norms are necessary to:
- i) ensure accountability - including allocating liability - from developers, owners, and implementers of AI systems;
 - ii) encourage continued investment in AI research and development; and

- iii) protect fundamental civil rights of all persons who are impacted by the adoption of the AI technology whether they are direct users or customers.
- b) **Recognize the usefulness of existing governance mechanisms and clarify their application where necessary.** Before new governance rules, procedures, or standards are developed for special-case aspects of AI, existing legal frameworks, especially those for computer-implemented, complex, and enabling technologies (with which AI shares several commonalities), should be reviewed for analogous models from which to develop a solution.
- c) **Prioritize international cooperation for ethical, trustworthy AI systems.** AI systems have pervasive and transformative global implications. The U.S. should proactively identify and address how these transformations will affect international relationships, as well as potential international governance frameworks. Broadly, the U.S. should prioritize and lead global cooperation, including with developing countries, across all fora to ensure that a comprehensive and interoperable governance scheme or entity is developed and implemented. The goal should be to align and harmonize best practices to ensure the safe and ethical development and use of AI systems and codify these best practices into international governance instruments.

This statement was developed by the IEEE-USA Artificial Intelligence Policy Committee and represents the considered judgment of a group of U.S. IEEE members with expertise in the subject field. IEEE-USA advances the public good and promotes the careers and public policy interests of the nearly 150,000 engineering, computing, and allied professionals who are U.S. members of the IEEE. The positions taken by IEEE-USA do not necessarily reflect the views of IEEE or its other organizational units.