

HIGHLIGHTS OF OECD AI REVISED RECOMENDATION BY THE COUNCIL ON ARTIFICIAL INTELLIGENCE

- Meeting of the Council at Ministerial level 2-3rd May 2024
 - Majiuzu Daniel Moses

FACTS

- The OECD AI Principles were first adopted in 2019
- It is the first intergovernmental standard on AI
- The 2024 Ministerial Council meeting adopted revisions in the landmark OECD Principles on AI.
- Currently 47 entities have adopted the revised principles including the EU.
- General scope ensures applicability of to AI development globally.

Definition of Terms

i. AI Systems

- A machine-based system that, for the explicit or implicit objectives, infers from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

ii. AI Stakeholders

Encompass all organizations and individuals involved in, or affected by AI systems, directly or indirectly.

iii. AI Lifecycle

- Involves several phases that include to: plan and design; collect and process data; build model(s) and/or adapt existing model(s) to specific tasks; test, evaluate, verify and validate; make available for use/deploy; operate and monitor; retire and decommission.
- The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase.

iv. AI Actors

- These are those who play active role in the AI system lifecycle, including organizations and individuals that deploy or operate AI.

v. AI Knowledge

- Refers to the skills and resources, such as data, code, algorithms, models, research, know-how, training programmes, governance, processes, and best practices required to understand and participate in the AI system lifecycle

Recommendations for Responsible stewardship of AI

1. Inclusive Growth sustainable development and well-being

Stakeholders to proactively pursue trustworthy AI beneficial to people and the planet such as:

- ✓ Augmenting human capabilities and enhancing creativity
- ✓ Advancing inclusion of underrepresented groups
- ✓ Reducing economic, social, gender and other inequalities
- ✓ Protecting natural environments for sustainable development

2. Respect for the rule of law, human rights and democratic values

- a. All actors should respect the rule of law, human rights, human-centered values throughout the AI system lifecycle. These include: non-discrimination, equality, privacy and data protection, diversity, fairness, social justice, int'l recognized labour rights, disinformation and misinformation.
- b. Need for mechanism and safeguards by AI actors such as human agency and oversight to address risks arising from intended purpose, intentional misuse and unintentional misuse.

3. Transparency and explainability

AI actors to ensure responsible disclosure regarding AI systems to:

- i. Foster general understanding of AI,
- ii. Make stakeholders aware of their interactions with AI,
- iii. Provide plain information on the sources of data/input, factors, processes and logic for prediction,
- iv. Enable those adversely affected by an AI system challenge its output.

4. Robustness, security and safety

- AI systems should be robust, safe and secure throughout their lifecycle.
- Mechanisms to override, repair or decommission and AI system should it exhibit undesired behavior(likely to cause risk) should be in place.
- Mechanisms to bolster information integrity while respecting freedom of expression should be in place where technically feasible.

5. Accountability

- AI actors should be accountable for the proper functioning of AI systems.
- AI actors should ensure traceability of datasets processes and decisions made during an AI system's lifecycle to enable analysis of its output.
- AI actors should on an ongoing basis, adopt a systemic risk management approach during each phase of the AI system lifecycle.

National Policies and Int'l co-operation for trustworthy AI

A. Investing in AI research and development

- ❑ Government should undertake long-term investment as well as encourage private investment in research and development to promote innovation in trustworthy AI.
- ❑ Government should embark on public investment as well as encourage private investment in open-source tools and datasets that respect privacy and data protection standards.

B. Fostering an inclusive AI ecosystem

□ Government should create dynamic and interoperable digital ecosystem for trustworthy AI which includes:

- Data
- AI technologies
- computational and connectivity infrastructure
- Sharing of AI knowledge

C. Shaping an enabling interoperable governance and policy environment for AI

- Government should support transitioning from research and development stage to the deployment and operation stage for trustworthy AI.
- Outcome based approach to be adopted
- Ensure that policies and regulatory frameworks are adaptable to encourage innovation and competition for trustworthy AI.

D. Building human capacity

- ❖ Government should closely work with stakeholders to integrate the world of work and the society with AI interaction.
- ❖ Ensure a fair transition into AI for workers through social dialogue, training, support and access to new opportunities.
- ❖ Work closely with stakeholders to promote responsible use of AI at work.

E. International Co-operation for trustworthy AI

- Governments(both developed and developing countries) should synergize to advance trustworthy AI.
- Governments should collaborate in the OECD and other global and regional for a to advance and share AI knowledge.
- Governments should promote development of global technical standards.
- Development of int'l comparable indicators to measure AI research, development and deployment.

Instruction to Digital Policy Committee

The Digital Policy Committee through its Working Party on AI Governance is instructed to:

- Continue its work on AI building on this recommendation and to further develop measurement framework for evidence-based AI policies;
- Develop further practical guidelines on the implementation of these recommendation;
- Provide a forum for exchanging information on AI policy;
- Report to Council on the implementation, dissemination and continued relevance of this recommendation no later than 5 years after this revision and every 10 years thereafter.

Conclusion

These revised recommendations by the OECD is in recognition of the rapidly evolving and dynamic AI landscape and the understanding that AI has far-reaching and global implications transforming societies, economic sectors and the world of work as well as the need to provide a comprehensive and human-centric framework for the responsible and trustworthy stewardship of AI. The recognition that trust is a key enabler of digital transformation, makes it necessary that trustworthiness of AI systems is a key factor for the adoption of AI.