What are the OECD Principles on AI?

Last update: 2 March 2020

Artificial intelligence is still in its early days and policymakers are still finding their feet. To what extent can they, and should they, encourage this powerful new technology, and how can they address any risks? The OECD Principles on AI can help. They promote artificial intelligence (AI) that is innovative and trustworthy and that respects human rights and democratic values.

They were adopted by OECD member countries when they approved the OECD Council Recommendation on Artificial Intelligence on 22 May 2019. As well as OECD member countries, Argentina, Brazil, Colombia, Costa Rica, Peru and Romania have also adhered to the AI Principles.

On 9 June 2019, the G20 adopted human-centred AI Principles that draw from the OECD AI Principles.

The five OECD AI Principles

The council recommendation identifies five complementary values-based principles for the responsible stewardship of trustworthy AI:

1. AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.

2. AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards–for example, enabling human intervention where necessary–to ensure a fair and just society.

3. There should be transparency and responsible disclosure around AI systems to ensure that people understand AI-based outcomes and can challenge them.

4. AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed.

5. Organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.

What can governments do?

Consistent with these value-based principles, the OECD also provides five recommendations to governments:

1. Facilitate public and private investment in research & development to spur innovation in trustworthy AI.

2. Foster accessible AI ecosystems with digital infrastructure and technologies and mechanisms to share data and knowledge.

3. Ensure a policy environment that will open the way to deployment of trustworthy AI systems.

4. Empower people with the skills for AI and support workers for a fair transition.

5. Co-operate across borders and sectors to progress on responsible stewardship of trustworthy AI.

What's next?

A future AI Policy Observatory will provide evidence and guidance on AI metrics, policies and practices to help implement the principles, and constitute a hub to facilitate dialogue and share best practices on AI policies. For more on AI see https://www.oecd.org/going-digital/ai/

References

See the Recommendation here: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449