



# AI GOVERNANCE AND THE ROLE OF THE OECD

Jerry Sheehan – Director, Directorate for Science, Technology and Innovation (STI)

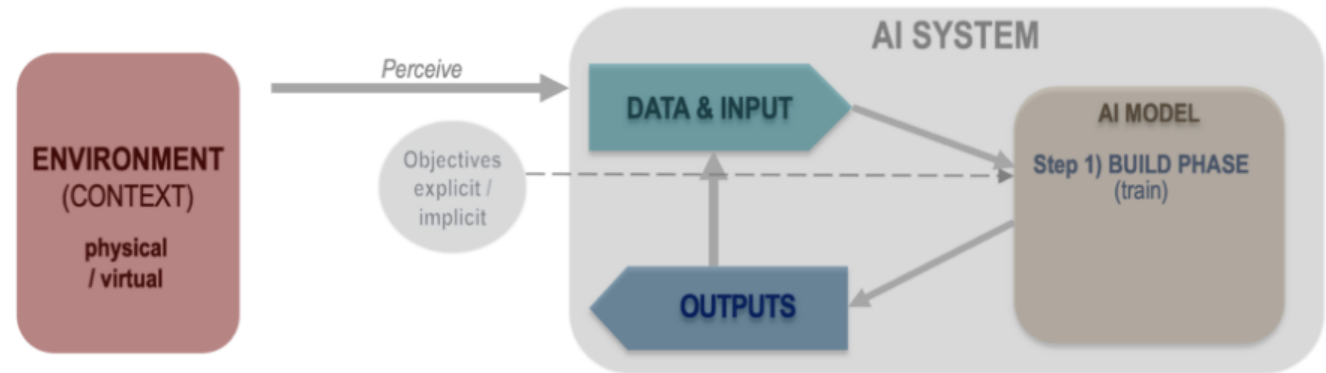
Karine Perset – Head, OECD.AI Policy Observatory (STI/DEP)

4 April 2024

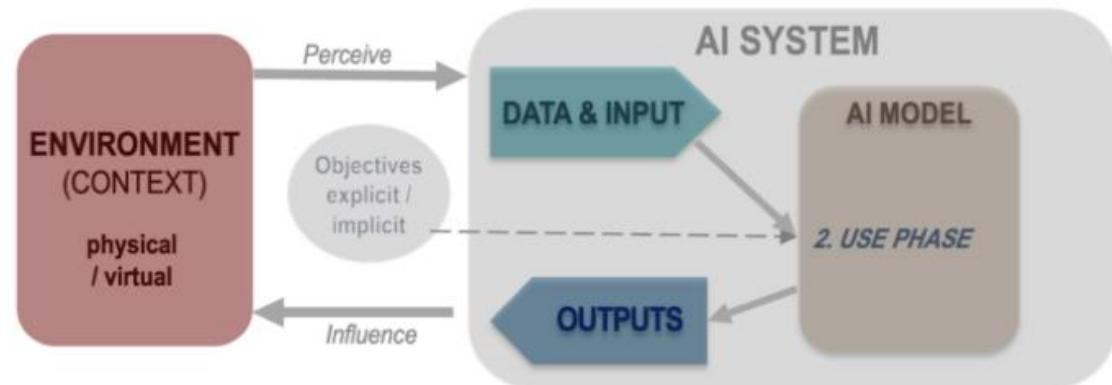
# What is Artificial Intelligence?

In November 2023, the OECD Council adopted an updated definition of an “AI system”:

## BUILDING AI SYSTEM



## USING AI SYSTEM



An **AI system** is:

“A machine-based system that, for explicit or implicit **objectives**,

**infers**, from the **input** it receives,

how to **generate outputs** such as predictions, content, recommendations, or decisions

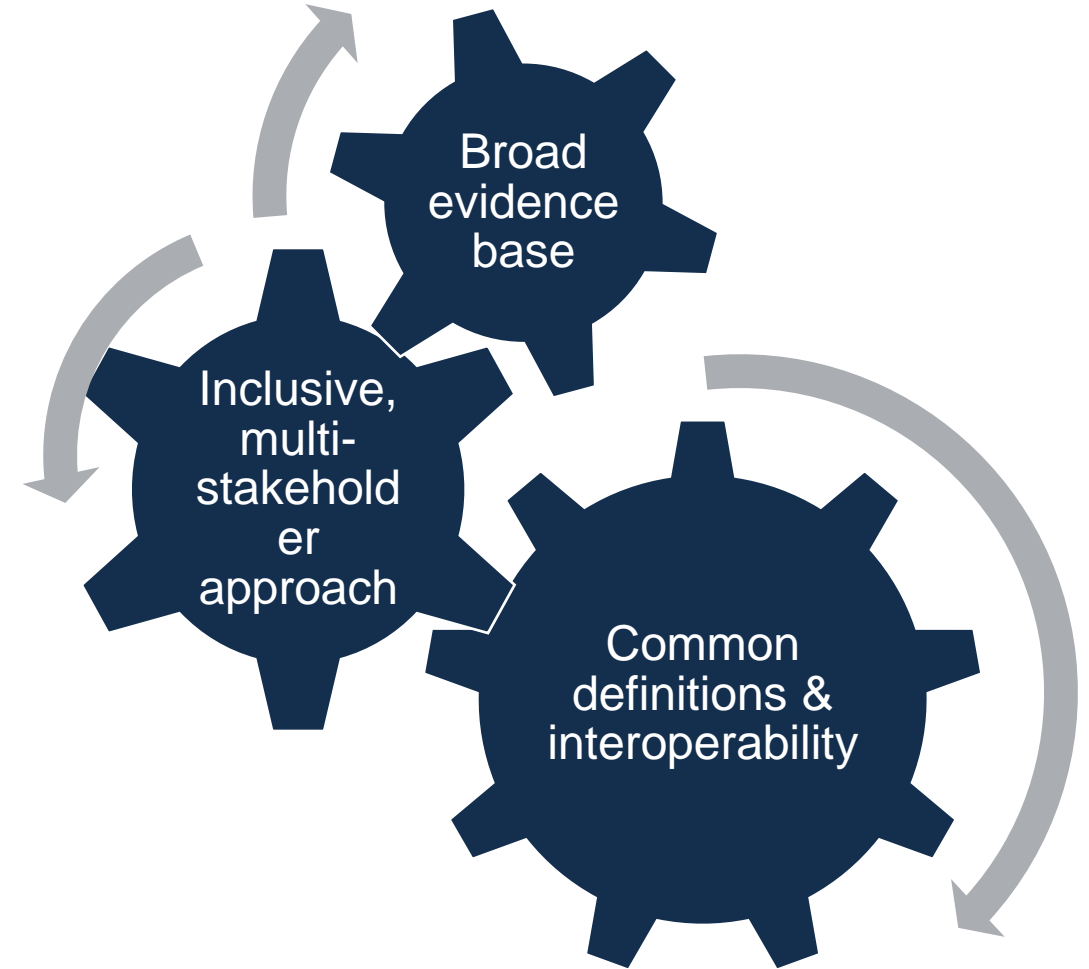
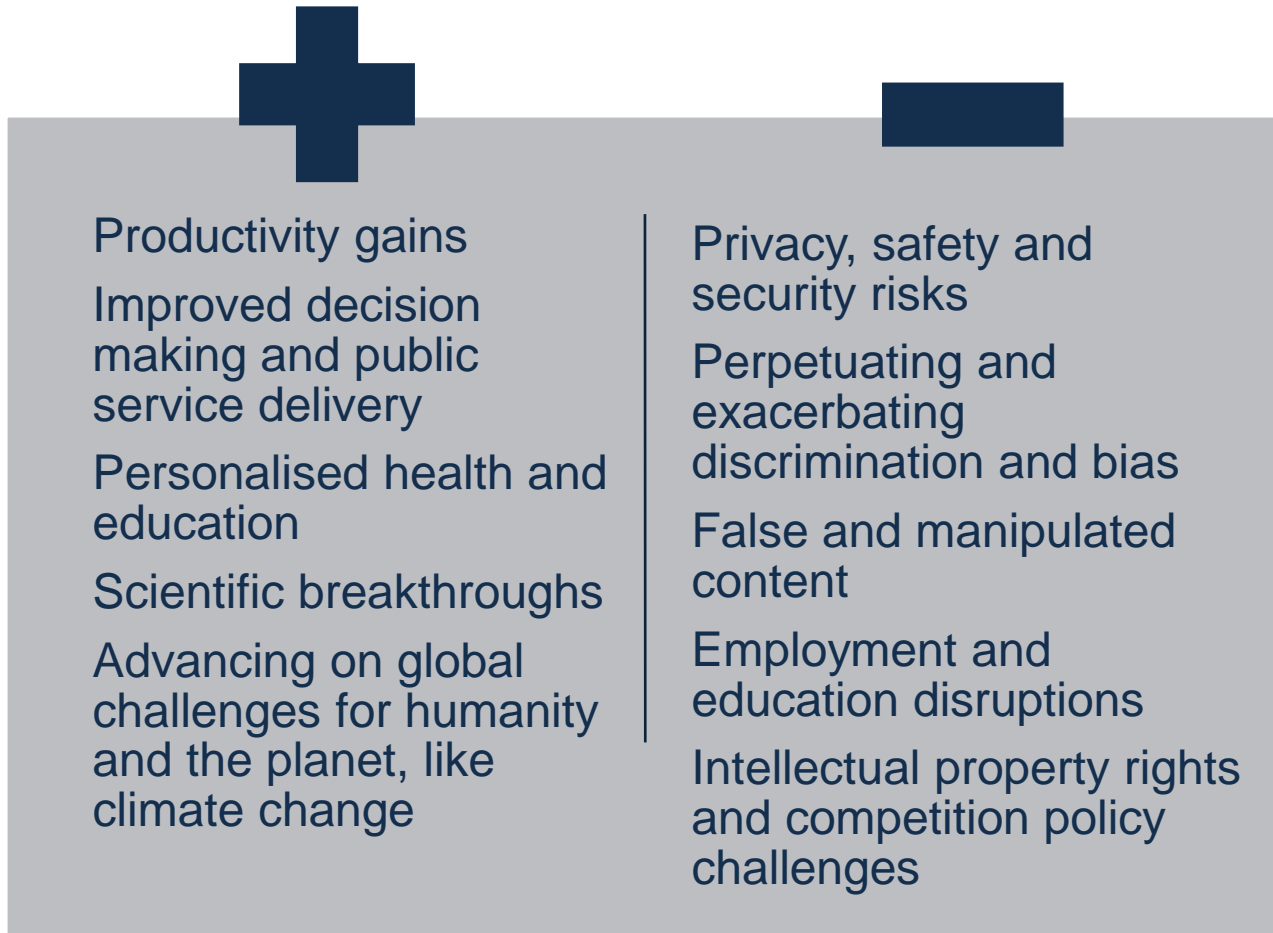
that can **influence** physical or virtual **environments**.

Different AI systems vary in their levels of **autonomy** and **adaptiveness** after deployment.”



# AI brings tremendous opportunities, but also risks

AI requires effective governance globally



2016

- G7 ICT Ministerial Meeting in Japan
- OECD Technology Foresight Forum on AI

2019

- **OECD Council Recommendation on AI**
- Report on “Artificial Intelligence in Society”
- **OECD Global Parliamentary group on AI**

2020

- OECD.AI Policy Observatory
- OECD Network of Experts (ONE.AI)

2022

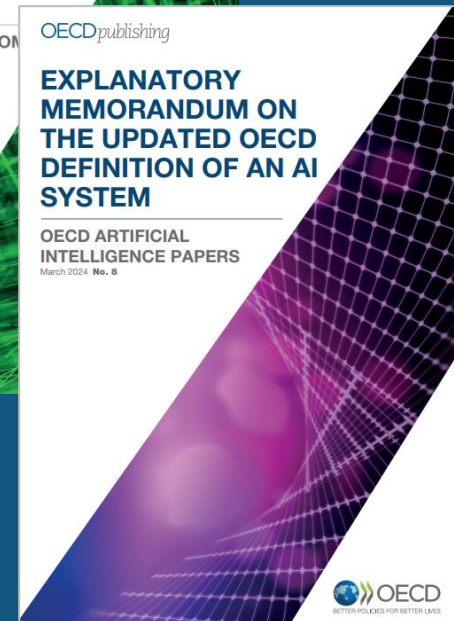
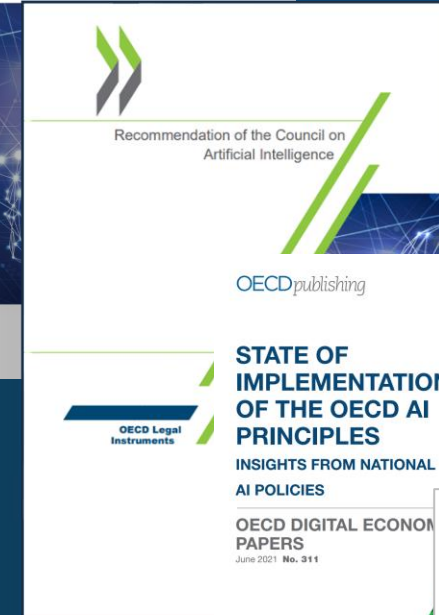
- OECD Working Party on AI Governance
- OECD Framework for the Classification of AI Systems

2023

- Report on "Advancing accountability in AI"
- Common guideposts to promote interoperability in AI risk management
- Revised definition of an “AI system”






2024

- Revising OECD Council Recommendation on AI



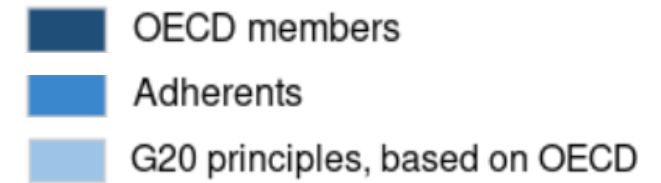
# The OECD AI Principles

## 5 values-based principles

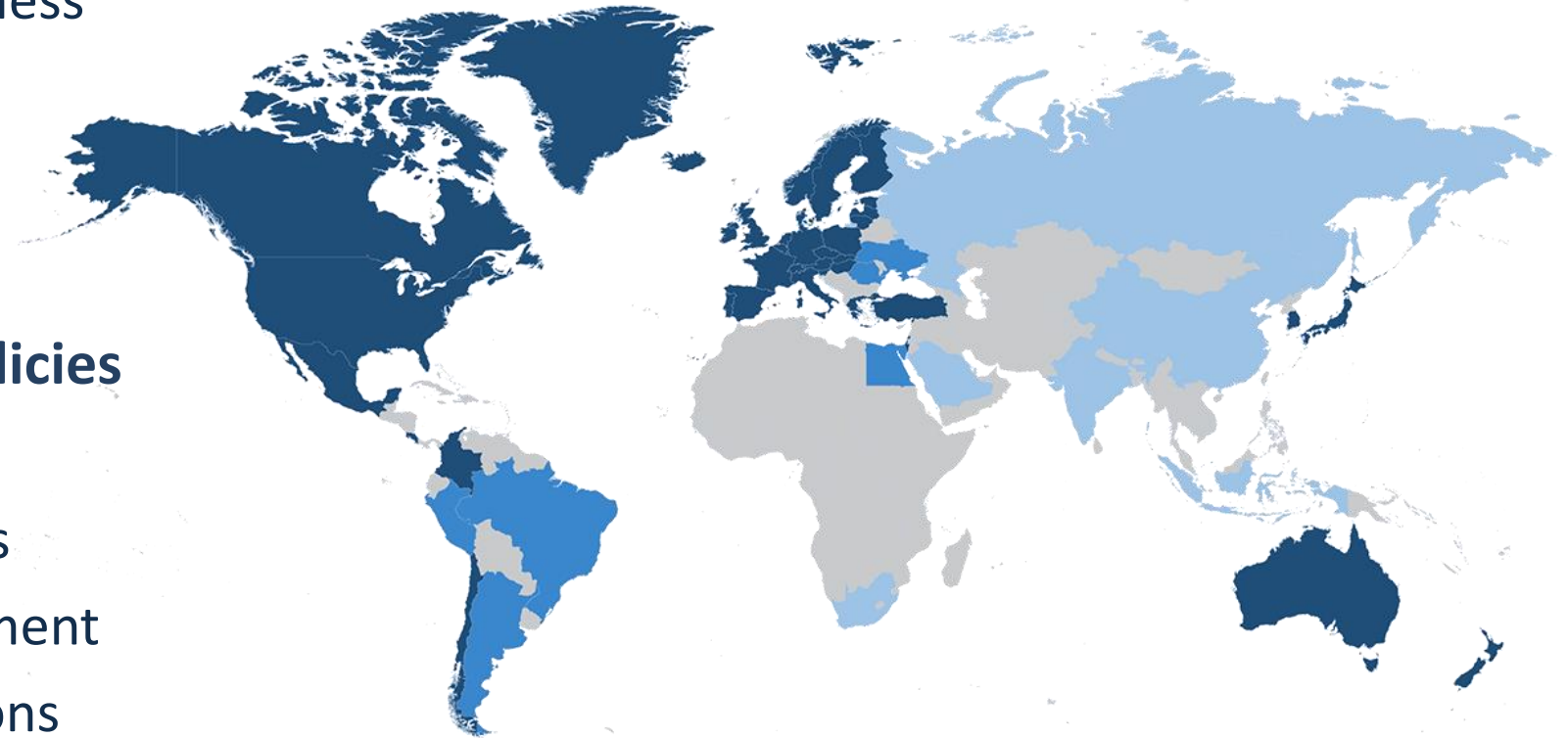
-  Benefit People & Planet
-  Human rights, values & fairness
-  Transparent & explainable
-  Robust, secure & safe
-  Accountable

## 5 principles for national policies

-  AI research & development
-  Data, compute, technologies
-  Policy & regulatory environment
-  Jobs & skills, labour transitions
-  International cooperation & measurement



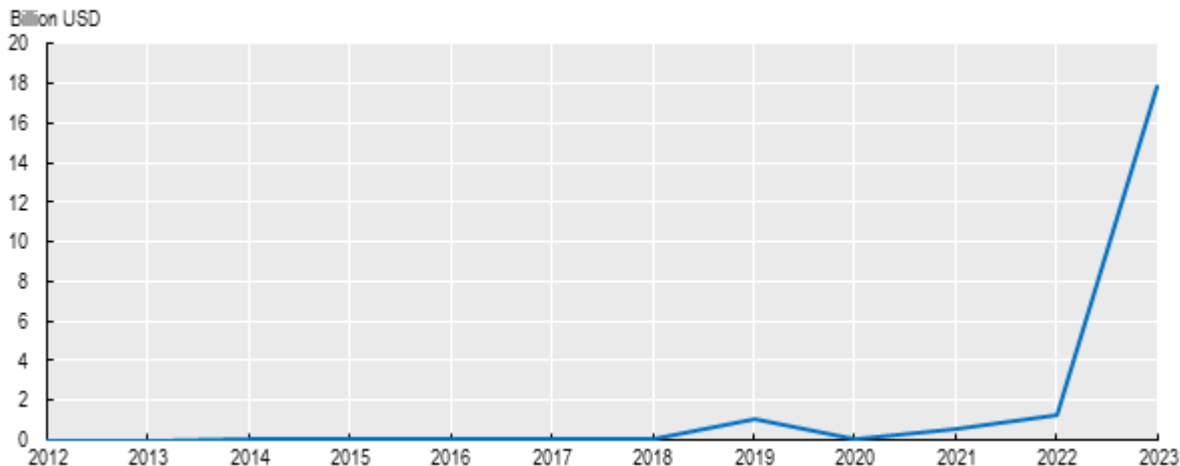
\*Singapore is an adherent



# What has changed in recent years?

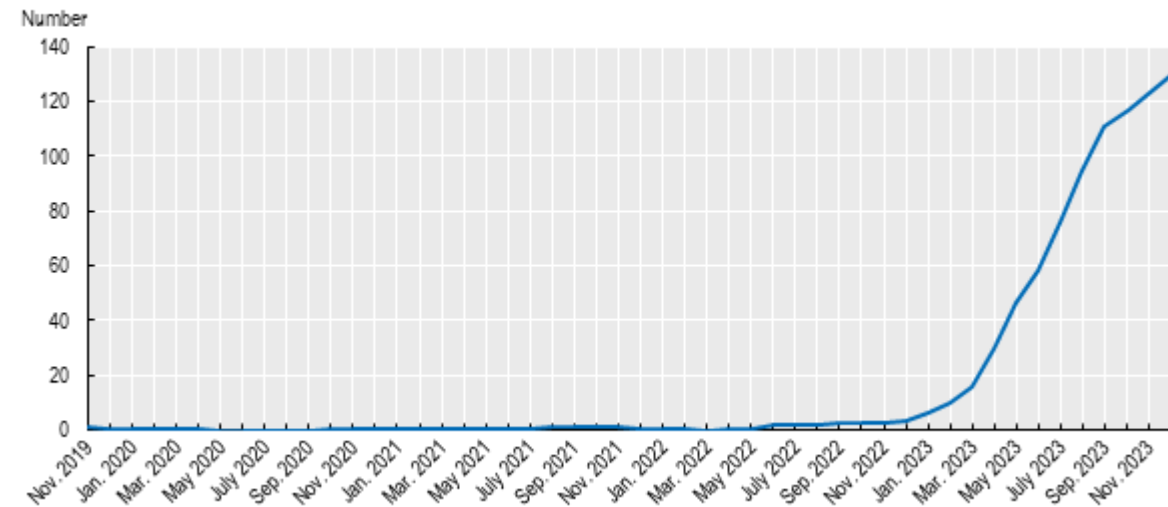
- Development, deployment and mass adoption of generative AI.
- Generative AI investments have grown rapidly, and so has the number of incidents related to the technology

Venture capital (VC) investments in generative AI start-ups have boomed since 2022



Source: OECD.AI (2024), using data from Preqin. Available at: [www.oecd.ai/en/data?selectedArea=investments-in-ai-and-data](http://www.oecd.ai/en/data?selectedArea=investments-in-ai-and-data)

Generative AI-related incidents and hazards have increased steeply since 2022



Source: OECD.AI (2024), AI Incidents Monitor (AIM), using data from Event Registry. Available at: [www.oecd.ai/incidents](http://www.oecd.ai/incidents) for more information.



[Home](#) > [AI Incidents Monitor](#)


With the support of



# OECD AI Incidents Monitor (AIM)



The OECD AI Incidents Monitor (AIM) offers an accessible starting point for comprehending the landscape of AI-related challenges. This initiative assumes significance in the context of AI's rapid proliferation across various domains. Whether you're a policymaker, a practitioner, or a concerned observer, the AIM equips you with insights into the incidents and hazards that shape the AI arena. As the AIM evolves, it aspires to foster a better understanding of AI incidents' multifaceted nature, empowering stakeholders to make informed decisions for the responsible development and deployment of AI technologies.

[> Please see methodology and disclosures here.](#) Enter a concept or a keyword[Search incidents](#)[Reset](#)

## ADVANCED SEARCH OPTIONS ▴

Date range:

All time



Country

Select country

Industry

Mob Select industry

AI principle

Select AI principle

Severity

Select severity

Harm type

Select harm type

Affected stakeholders

Select affected stakeholders

Type of search

Any of the concepts/keywords (OR) ▾







elections




Search incidents


Reset



### 'Terrifying prospect': Top AI expert warns of threat to democracy

 2024-02-22

 2 articles

 Ireland

Rapid recent advances in artificial intelligence (AI) systems are leading to increasing concern that the output of the technology will be misused during forthcoming elections. In 2024, approximately half the world's adult population will be asked to vote. Over the last twelve months, several AI companies have released software and tools which make it possible to easily create images, audio and video which are almost indistinguishable from authentic versions of the same content. As a result, r...

- AI principles:
- Democracy & human autonomy
- Transparency & explainability
- Industries:
- Digital security
- Affected stakeholders:
- General public
- Government
- Harm types:
- Public interest
- Severity:
- Non-physical harm

► Why's our monitor labelling this an incident?



### Home Secretary warns deepfakes could 'hijack' British elections

 2024-02-25

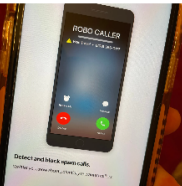
 6 articles

 United Kingdom

Deepfakes generated by AI provide the 'perfect storm' for those looking to hijack the upcoming general election, the Home Secretary has warned. James Cleverley said the era of these technologies misleading and disrupting important democratic decisions was 'already in play', The Times reported. It comes as Mr Cleverly is reportedly due to meet with Silicon Valley bosses to urge greater action from large technology companies to protect democracy. Earlier this month some of the companies signed ...

- AI principles:
- Democracy & human autonomy
- Industries:
- Digital security
- Government, security and defence
- Affected stakeholders:
- General public
- Government
- Consumers
- Harm types:
- Public interest
- Psychological

► Why's our monitor labelling this an incident?



### Biden deepfake in NH sounds alarm on AI in elections: New law proposed

 2024-02-06

 2 articles

 United States

HANCOCK -- At 5:35 p.m. on Jan. 21, just two days before New Hampshire's first-in-the-nation primary, Gail Huntley received a phone call with President Joe Biden telling her not to vote. "I didn't think I was hearing him right," said Huntley, 73, of her immediate reaction. A Democrat from Hancock, she was planning to write in Biden's name on the New Hampshire primary ballot. "It's important that you save your vote for the November election," the voice on the other side of the phone said. "Voti...

- AI principles:
- Privacy & data governance
- Robustness & digital security
- Democracy & human autonomy
- Respect of human rights
- Industries:
- Digital security
- Government, security and defence
- Affected stakeholders:
- General public
- Harm types:
- Psychological
- Reputational



### Legally Banned Deepfake Campaign Contents Rampant ahead of April's Elections

 2024-02-19

 8 articles

 Korea, Republic of

Legally banned deepfake contents are on the rise with around 50 days left until the April 10 general elections. According to the National Election Commission(NEC) on Monday, there have been 129 deepfake contents targeting voters in violation of the public election law between January 29 and February 16. The NEC said most of the contents have been deleted, while measures are being taken for others. AI-based fabricated contents or images have emerged as a major potential threat to democracy. Thi...

- AI principles:
- Democracy & human autonomy
- Industries:
- Digital security
- Government, security and defence
- Affected stakeholders:
- General public
- Government
- Business
- Harm types:
- Public interest
- Human rights
- Reputational
- Severity:
- Non-physical harm

► Why's our monitor labelling this an incident?



### AI 'resurrects' long dead dictator in murky new era of deepfake electioneering. | CNN

 2024-02-12

 7 articles

 Indonesia

An AI-generated deepfake of the late dictator Suharto has sparked debate about using AI technology for political gain. CNN -- A once-feared army general, who ruled Indonesia with an iron fist for more than three decades, has a message for voters ahead of upcoming elections - from beyond the grave. "I am Suharto, the second president of Indonesia," the former general says in a three-minute video that has racked up more than 4.7 million views on X and spread to TikTok, Facebook and YouTube. Wh...

- AI principles:
- Democracy & human autonomy
- Industries:
- Government, security and defence
- Affected stakeholders:
- General public
- Consumers
- Harm types:
- Reputational
- Psychological
- Severity:
- Non-physical harm



### Female MPs could face threat of deepfake porn at general election

 2024-01-23

 2 articles

 United Kingdom

Female MPs could be targeted by deepfake pornography, the elections chief has suggested. John Pullinger, the chairman of the Electoral Commission, warned that a rise in misinformation generated by artificial intelligence (AI) could "block out the real campaign" in the run-up to the election. Mr Pullinger said that he was concerned by the rise in deepfake porn "which inevitably is going to be much more targeted towards female candidates". He told the Financial Times that elected politicians in...

- AI principles:
- Respect of human rights
- Robustness & digital security
- Democracy & human autonomy
- Privacy & data governance
- Industries:
- Digital security
- Affected stakeholders:
- General public
- Consumers
- Harm types:
- Reputational
- Public interest
- Psychological



# The OECD.AI Policy Observatory


## National AI policies & strategies

This section provides a live repository of over 1000 AI policy initiatives from 69 countries, territories and the EU. Click on a country/territory, a policy instrument or a group targeted by the policy.

Countries & territories

Policy instruments

Target Groups

 Search for a specific das

Download all AI policies

 African Union	 Costa Rica	 Iceland	 Luxembourg	 Romania	 Tunisia
 Argentina	 Croatia	 India	 Malta	 Rwanda	 Türkiye
 Armenia	 Cyprus 	 Indonesia	 Mauritius	 Saudi Arabia	 Uganda
 Australia	 Czechia	 Ireland	 Mexico	 Serbia	 Ukraine
 Austria	 Denmark	 Israel	 Morocco	 Singapore	 United Arab Emirates
 Belgium	 Egypt	 Italy	 Netherlands	 Slovakia	 United Kingdom
 Brazil	 Estonia	 Japan	 New Zealand	 Slovenia	 United States
 Bulgaria	 Finland	 Kazakhstan	 Nigeria	 South Africa	 Uruguay
 Canada	 France	 Kenya	 Norway	 Spain	 Uzbekistan
 Chile	 Germany	 Korea	 Peru	 Sweden	 Viet Nam
 China	 Greece	 Latvia	 Poland	 Switzerland	 European Union
 Colombia	 Hungary	 Lithuania	 Portugal	 Thailand	





# Select emerging AI-specific regulatory frameworks

*This table is draft and for discussion purposes only. It is illustrative and not an exhaustive list of AI-specific regulations.*

Country	Legislation and regulation	Standards and toolkits	Enforcement mechanisms	Principles
Canada	<ul style="list-style-type: none"><li>Directive on Automated Decision-Making (2019)</li><li>Bill C-27, Digital Charter Implementation Act, including AI and Data Act (AIDA) (2022, draft)</li></ul>	<ul style="list-style-type: none"><li>CAN-ASC-6.2: Accessible and Equitable Artificial Intelligence Systems (2023, draft)</li></ul>	<ul style="list-style-type: none"><li>AI and Data Commissioner (forthcoming)</li><li>Courts</li></ul>	<ul style="list-style-type: none"><li>Canada's Digital Charter (2019)</li></ul>
United States	<ul style="list-style-type: none"><li>Federal Trade Commission Act, Section 5 prohibition on unfair or deceptive acts or practices in or affecting commerce (1914)</li><li>Algorithmic Accountability Act (US AAA) (2022, draft)</li><li>Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (2023)</li></ul>	<ul style="list-style-type: none"><li>NIST AI Risk Management Framework (2023)</li></ul>	<ul style="list-style-type: none"><li>Authorities at the national and local levels regulate context and sector-specific uses of AI</li><li>Courts</li></ul>	<ul style="list-style-type: none"><li>Blueprint for an AI Bill of Rights (2023)</li></ul>
United Kingdom	<ul style="list-style-type: none"><li>Data Protection and Digital Information Bill (2023, draft)</li><li>Online Safety Bill (2022, draft)</li></ul>	<ul style="list-style-type: none"><li>AI Standard Hub (forthcoming)</li><li>Algorithmic Transparency Standard (Central Digital Data Office, 2021)</li></ul>	<ul style="list-style-type: none"><li>Sectoral and cross-sectoral authorities; exchange for policy coherence by authorise (CMA, Ofcom, FCA) under the Digital Regulation Cooperation Forum</li><li>Courts</li></ul>	<ul style="list-style-type: none"><li>Establishing a pro-innovation approach to AI regulation (2023)</li></ul>
European Union	<ul style="list-style-type: none"><li>Proposed EU AI Act (2021, draft)</li><li>Updates to the EU Product Liability Directive (2022, draft)</li><li>AI Liability Directive (2022, draft)</li></ul>	<ul style="list-style-type: none"><li>CEN/CENELEC standards for AI and related data (forthcoming)</li></ul>	<ul style="list-style-type: none"><li>EU AI board for oversight (forthcoming)</li><li>National authorities in charge of administrative sanctions</li><li>Civil and criminal sanctions at the national level (courts)</li></ul>	<ul style="list-style-type: none"><li>Ethics guidelines on artificial intelligence (2018)</li></ul>
Brazil	<ul style="list-style-type: none"><li>Report and proposed substitute text for draft bills 5051/2019, 21/2020 and 872/2021 (2022, draft)</li><li>Bill 705 on the compatibility of AI use in the public sector with ESG practices (2022, draft)</li></ul>	<ul style="list-style-type: none"><li>Incorporation of international standards National standards by the Brazilian Association of Technical Norms (ABNT)</li></ul>	<ul style="list-style-type: none"><li>Designated competent authority (to be determined) in charge of oversight, administrative sanctions and regulatory sandboxes</li><li>Courts</li></ul>	<ul style="list-style-type: none"><li>Art. 3 of the proposed substitute text for draft bills 5051/2019, 21/2020 and 872/2021 (2022, draft)</li></ul>
China	<ul style="list-style-type: none"><li>Chinese Internet Information Service Algorithmic Recommendation Provisions (2021)</li><li>Opinion on Strengthening the Ethics and Governance of Science and Technology (2022)</li></ul>	<ul style="list-style-type: none"><li>National Standards for Autonomous Vehicle Testing (2018)</li></ul>	<ul style="list-style-type: none"><li>Cyberspace Administration of China, among other regulators competent over specific matters (telecommunications, public security, market regulation, competition policy)</li><li>National Data Administration (2023, forthcoming)</li><li>Courts</li></ul>	<ul style="list-style-type: none"><li>New Generation AI Ethics Specifications (2019)</li><li>New Generation AI Code of Ethics (2021)</li><li>White Paper on Trustworthy AI (2021)</li><li>Internet Information Service Algorithmic Recommendation Management Provisions (2021)</li></ul>
IGOs	<ul style="list-style-type: none"><li>Council of Europe Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (2023, draft)</li></ul>	<ul style="list-style-type: none"><li>ISO 31000 Risk management (2009, 2018)</li><li>ISO/IEC 23053:2022 Framework for AI Systems Using Machine Learning (ML) (2022)</li></ul>	<ul style="list-style-type: none"><li>Each signatory State to set in place measures and safeguards to ensure accountability and redress at the national level (Chapter V of the Convention)</li></ul>	<ul style="list-style-type: none"><li>OECD Recommendation of the Council on AI (2019)</li><li>UNESCO Recommendation on the Ethics of AI (2021)</li></ul>

# Risks of fragmentation in global AI policies?

**Today:** Emergence of AI regulatory frameworks, standards, and risk management frameworks in different countries and regions.



Addressing AI risks



Legal complexities & high implementation costs for businesses

**The OECD promotes alignment & interoperability between AI frameworks to create certainty for policy makers, AI users, AI developers, and businesses**



*Thank you!*

For more information please visit  
[www.oecd.ai](http://www.oecd.ai)