

OECD Input to the United Nations High-Level Advisory Body on Artificial Intelligence

1. Background

The Office of the UN Secretary-General's Envoy on Technology (OSET), in its capacity as the Secretariat of the UN [High-level Advisory Body on Artificial Intelligence](#) (UN AI Advisory Board) is seeking input, including from the OECD, on the [Interim Report: Governing AI for Humanity](#). This input will contribute to finalising the report, which will serve as a key basis for the [UN Summit of the Future](#) in September 2024.

Consequently, the OECD Working Party on Artificial Intelligence Governance (AIGO) and its Secretariat have put together this input, in consultation with other parts of the OECD Secretariat, summarising relevant past and ongoing work strands of the OECD, addressing in particular regarding '*functions for international governance of artificial intelligence*' described in the report (Table 1) and selected '*key questions for further discussion in the next phase of work*' of the UN AI Advisory Board. This input was considered at the 7th and 8th meeting of AIGO on 29 February 2024 and 26 March 2024, respectively. It is submitted pursuant to OSET's request for input by 31 March 2024.

This document outlines some of the ongoing activities of the OECD in the realm of AI and how its role in this area could be further developed in cooperation with the UN. The OECD, including AIGO, looks forward to continuing and deepening engagement with the United Nations moving forward.

2. Comments on the overall report

The UN has undisputed expertise in delivering capacity building, AI for good and social benefit, and promoting and protecting international human rights law. The OECD's primary function is as a forum for expert analysis and sharing of best practice, working with business, which can be conducted in a complementary manner to UN-led workstreams.

In terms of possible improvements to the report, its scope could perhaps be made more precise by, for example, specifically stating what "AI", "AI systems", and "AI governance" refer to in the context of the report, and might wish to leverage, as appropriate, the [Explanatory memorandum on the updated OECD definition of an AI system](#).

In addition, it might be helpful to further elaborate the section on "The AI governance landscape", by highlighting the different areas in which the governance frameworks related to AI that are already in place should be interconnected, for example distinguishing governance of technical infrastructures, procedures and practices, from governance of human behaviour.

A global AI governance framework should also build upon other existing domestic and international AI governance efforts by other institutions.

3. OECD.AI inputs on functions for international governance of artificial intelligence

Table 1. Summary table of subfunctions for international governance of artificial intelligence, and possible timeframes for realisation (from Interim Report, page 22) [with OECD comments]

Subfunction	Description	Category	Possible timeframe required to institutionalise proposed subfunction	OECD input
1. Scientific assessment	Prepare a public review of international, regional, and national AI policies at least every 6 months.	Research & Analysis	6-12 months	<p>The OECD Directorate for Science, Technology and Innovation (DSTI) has been maintaining the OECD.AI database of national and regional AI policies since 2018.</p> <ul style="list-style-type: none"> This is updated on a rolling basis and formally every 6 months by way of a survey to national contact points and serves to inform a now-annual review, the latest of which was published in October 2023. The survey has benefited from input from many parts of the OECD, including the Development Co-operation Directorate (on AI in development); the Centre for Entrepreneurship, SMEs, Regions and Cities (on AI and entrepreneurship); and the Public Governance Directorate (on AI in the public sector). The database already includes information on 70 jurisdictions including numerous OECD partner countries from Africa, Latin America, and Asia. It is being further expanded to include additional policies of countries participating in the African Union Commission's Working Group on AI. <p>The OECD Directorate for Employment, Labour and Social Affairs has been monitoring the impact of AI on society – including on labour markets and health.</p> <ul style="list-style-type: none"> These research efforts aim to help policy makers better understand the impact of AI on the labour market, including the various risks and benefits. <p>POSSIBLE ACTION:</p> <ul style="list-style-type: none"> The database and its activities could be extended to include member countries of the United Nations which are not currently included, and the database possibly branded to reflect such a tripartite co-operation (currently, the database is a joint project with the European Commission). Building on this pre-existing authoritative tool (the OECD national AI policies database) would help to minimise additional effort involved for member countries to provide relevant information on national AI policies. Reports could developed by different organisations in partnership or separately.

2. Horizon scanning	Prepare a horizon-scanning report that identifies risks that transcend borders and can potentially affect all jurisdictions.	Research & Analysis	6-12 months	<p>In 2023, OECD-DSTI, in cooperation with the OECD’s General Secretariat’s Foresight Unit created an OECD.AI Expert Group on AI Futures, of which the mandate is explicitly to conduct horizon scanning of AI-related risks that transcend borders as well as scenario building and solution identification.</p> <p>In the area of horizon scanning, it will be critical to ensure complementarity with existing international AI safety initiatives, notably the International AI Safety Report and Summits as credible options to fulfil some of the international governance functions proposed.</p> <p>POSSIBLE ACTION: This group could include a UN Secretariat representative to identify possible synergies.</p>
3. Risk classification	Assess existing and upcoming AI models on a risk scale of untenable, high-level, mid-level, and low to no risks.	Research & Analysis	6-12 months	<p>OECD-DSTI published a report early 2022 on the Classification of AI Systems, which could be helpful to defining such risk thresholds. The European Commission and several standards bodies including ISO, CEN-CENELEC and NIST are also working on risks scales. They are coordinating some of this work through the OECD.AI expert group on AI Risk and Accountability.</p> <p>Recent OECD publications on this topic include:</p> <ul style="list-style-type: none"> • Advancing accountability in AI: This report provides an overview of how risk management frameworks and the AI system lifecycle can promote trustworthy AI. • Common guideposts to promote interoperability in AI risk management: This report provides an analysis of the commonalities of AI risk management frameworks. It demonstrates that, while some elements may sometimes differ, all the risk management frameworks analysed follow a similar and sometimes functionally equivalent risk management process. <p>POSSIBLE ACTION: The OECD and the UN could for example partner on workshops on risk classification based on existing work (and collaboration with OHCHR by the OECD.AI Expert Group on Risk and Accountability) and supported by the OECD.AI network of experts.</p>
4. Access to benefits	Equitable access to technology and benefits of AI, accelerating achievement of the Sustainable Development Goals.	Enabling	12-24 months	<p>This area seems to be well aligned with the UN’s mandate and areas of expertise.</p> <p>OECD-DSTI has conducted analysis on measuring the environmental impact of AI as well as reviews examining the use of AI to advance particular sectors, such as in its report “The state of implementation of the OECD AI Principles four years on.” Sectoral work on AI has been conducted by different parts of the OECD without a specific focus on the SDGs, such as:</p> <ul style="list-style-type: none"> • Education • Employment • Health • Public governance, e.g. The Strategic and Responsible Use of AI in the Public Sector of Latin America and the Caribbean • Science • Finance, e.g. Business and Finance Outlook, 2021, Chapter 1 – Trends and policy frameworks for AI and Chapter 2 – AI in Finance • On-going and forthcoming OECD work on AI and environment include the OECD’s Horizontal Project “Net Zero+: Building Climate and Economic Resilience in a Changing World” (Net Zero+). <p>The UN could leverage these resources as appropriate.</p>

<p>5.Capacity building</p>	<p>Programs and resources to build AI technology and businesses as well as governance and promotional capacity among states.</p>	<p>Enabling</p>	<p>12-24 months</p>	<p>In 2021, OECD-DSTI jointly with the Inter-American Development Bank published a practical toolkit for using AI in the public sector entitled: Responsible use of AI for public policy: Data science toolkit.</p> <p>Several OECD reports analyse the public sector and business capacities necessary to build and adopt AI in-depth:</p> <ul style="list-style-type: none"> • The state of implementation of the OECD AI Principles four years on (2023) on how governments and innovators are testing innovative AI solutions in controlled environments • The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean (2022) • Chapters on “New forms of accountability for a new era of government” in Global trends in government-innovation (2023) • State of implementation of the OECD AI Principles (2021) • Artificial intelligence, digital technology and advanced production (2020) • Hello, World! Artificial intelligence and its use in the public sector (2019) • Review of national policy initiatives in support of digital and AI-driven innovation (2019) • State of the art in the use of emerging technologies in the public sector (2019) • AI in Society (2019) <p>Through thematic and country-specific analyses (notably on Governing with AI, Germany and Egypt, both forthcoming), the OECD is supporting countries in enhancing their capacities in strategic AI policy development.</p> <p>The OECD also facilitates capacity-building on a regional policy level, including through a Global AI Governance Dialogue with the African Union and the OECD-MENA Working Group on Open and Innovative Government.</p> <p>The OECD has emphasised the need to ensure the regulatory and governance approaches continue to be relevant in the Recommendation for Agile Regulatory Governance to Harness Innovation, which is relevant in the rapidly changing AI landscape.</p> <p>POSSIBLE ACTION: Targeted policy advice and capacity-building measures could be developed in partnership between OECD and UN and undertaken to support governments’ capacity to govern AI and govern with AI. These could leverage tools such as the OECD catalogue of tools and metrics of trustworthy AI (oeecd.ai/tools), and other relevant resources (oeecd.ai/resources and GlobalPolicy.ai/search).</p>
<p>6.Joint R&D</p>	<p>Establish the capacity to undertake collaborative research and development of AI to benefit those who don’t have access to AI tools or expertise.</p>	<p>Enabling</p>	<p>12-24 months</p>	<p>Several OECD reports and activities delve into the benefits of wider access to and expertise in AI tools, of collaborative research and development:</p> <ul style="list-style-type: none"> • Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research (2023) • Identifying and characterising AI adopters: A novel approach based on big data (2022) • AI diffusion in firms: what do we know and what does it mean for policy? (2021) • Knowledge co-creation in the 21st century (2021) • Building digital workforce capacity and skills for data-intensive science (2020) • Digital Innovation: Seizing Policy Opportunities (2019) <p>POSSIBLE ACTION: Policy advice could be developed in partnership between OECD and UN and good practices could be shared to support governments in providing access to AI tools and expertise (e.g. statistical capacity-building), as well as building an enabling environment for AI responsible and trustworthy use in the public sector (e.g. AI skills development in public workforce).</p>

<p>7. Inclusive participation</p>	<p>Ensure participation of all stakeholder groups and all countries and regions in collective governance, risk management and realization of opportunities; strive for innovative governance.</p>	<p>Governing</p>	<p>6-12 months</p>	<p>OECD analyses have raised the need for accountability and participatory risk management frameworks for AI, such as the report on Accountability in AI: Governing and managing risks as well as Common guideposts to promote interoperability in AI risk management, and discussed the challenges of effective governance of AI.</p> <p>Further analyses have explored in-depth innovative governance tools for AI, such as AI regulatory sandboxes.</p> <p>Participation of a wide variety of stakeholders in technology governance has been a focus across the work of OECD-DSTI (e.g. Engaging citizens in innovation policy or the Inclusive Innovation Policy Toolkit) as has been the effects of technology on particular societal groups (e.g. The effects of AI on the working lives of women).</p> <p>Such issues have also been a focus of the OECD Public Governance Directorate (GOV), particularly through its inclusive approach to the design and delivery of digital and innovative public services and policies, publications on New ways of engaging citizens and residents (Chapter 4) and Surfacing ground-up insights and collective intelligence across borders (Chapter 1), Good practice principles on service design and delivery, and through its work on Innovative Citizen Participation and Civic Space.</p> <p>A report on Collective Action for Responsible AI in Health focusses on the need for collective action across governments and sectors. The subsequent brief AI in Health: Huge Potential, Huge Risks articulates the risks of inaction leading to expansion of current digital divides. These propose that collective action would foster trustworthiness of AI solutions (for the public and providers), enable human and technical capacity for AI, establish a common policy and standards environment for AI that works across borders, and evaluate the impacts of AI solutions to support its evolution.</p>
<p>8. Convening, international learning</p>	<p>Convene stakeholders regularly to consider AI policies across jurisdictions; building consensus on shared vocabulary and definitions; peer to peer learning.</p>	<p>Governing</p>	<p>6-12 months</p>	<p>AI legislation and regulation need definitions to establish a solid foundation and promote interoperability between jurisdictions. The OECD's definition of AI systems, updated in November 2023, is important to ongoing legislative work across several jurisdictions. OECD analysis has also provided frameworks and definitions for the AI system lifecycle, Classifying AI systems, AI risk management and developing a Framework for monitoring AI incidents.</p> <p>The OECD.AI network of experts, an informal group of AI experts provides AI-specific policy advice for the OECD's work on AI policy. It includes representatives from all stakeholder groups: government, business, academia, and civil society. Efforts are ongoing to broaden participation beyond member and partner economies and these could be further strengthened in cooperation with the UN.</p> <p>This network of experts is being expanded to include sector-specific considerations (e.g. employment, health care) to support practical implementation of AI practices that respects sector-specific characteristics while being grounded in overall AI leading practices. The OECD for example is planning to develop an Action Plan to Harness the Benefits and Address the Risks of Artificial Intelligence in the Labour Market.</p> <p>Governance considerations should recognise that sector-specific initiatives are underway and that such initiatives should be aligned with horizontal governance.</p> <p>The annual OECD International Conference of the Programme on AI in Work, Innovation, Productivity, and Skills (AI-WIPS) brings together leading voices from the policy, academic, business, technical and civil society communities to discuss how AI affects employment, skills, productivity, and innovation, and how policymaking can respond.</p>

<p>9. International coordination</p>	<p>Deconflicting work and building synergy across existing international bodies that continue to address AI.</p>	<p>Governing</p>	<p>6-12 months</p>	<p>In the framework of the GlobalPolicy.ai platform, the Council of Europe, the European Commission, European Union Agency for Fundamental Rights, Inter-American Development Bank, the OECD, the United Nations, UNESCO and the World Bank, cooperate based on their complementary mandates on AI to help policy makers and the public navigate the international AI governance landscape and access the necessary knowledge, tools, data, and best practices to inform AI policy development.</p> <p>The objectives are:</p> <ul style="list-style-type: none"> to provide information about each organisation’s work on AI, thereby helping policy makers and the wider public navigate the international AI governance landscape and access relevant resources; and to strengthen co-operation between the different organisations, leveraging each other’s work and making progress together towards common goals. <p>Key focus areas of collaboration of the group are:</p> <ul style="list-style-type: none"> Responsible development and use of trustworthy and ethical AI AI aligned with human rights and democracy AI for the Sustainable Development Goals (SDGs) <p>The group meets every few months to discuss how the participating organisations can strengthen cooperation and leverage each other’s work to achieve common goals for trustworthy, human-centred AI and explore concrete collaborative projects, such as joint policy reports and events. An initiative that originated with globalpolicy.AI is a global challenge to build trust in the age of generative AI.</p> <p>The OECD has also supported the Global Partnership on AI (GPAI) by hosting its Secretariat.</p>
<p>10. Policy harmonization; norm alignment</p>	<p>Surfacing best practices for norms and rules, including for risk mitigation and economic growth. Align, leverage, and include, soft and hard law, standards, methods, and frameworks developed at the regional, national, and industry level to support interoperability.</p>	<p>Governing</p>	<p>12-24 months</p>	<p>The OECD AI Principles were the first intergovernmental standard on AI. Adopted by OECD member and non-OECD countries in 2019, they also constitute the basis for the G20 AI Principles. They aim to foster innovation and trust in AI by promoting the responsible stewardship of trustworthy AI while ensuring respect for human rights and democratic values. The first five of these principles are value-based principles related to ethical and safety questions for AI to be trustworthy & human-centric and these apply to all actors. The five other lay out key actions that policy makers need to implement in order to foster an AI ecosystem that can thrive on the long-term and benefit societies. As well, the OECD AI Principles can help guide the development of AI policies, regulations, and standards at the regional, national, and industry level and in turn support their alignment through a common set of principles.</p> <p>The Catalogue of tools & metrics for trustworthy AI is a centralised resource offers diverse tools, mechanisms and practices to help all AI practitioners to ensure their systems’ trustworthiness. It was developed by the OECD in partnership with NIST, the European Commission, the UK AI Standards Hub and the Partnership on AI.</p> <p>Additionally, the abovementioned OECD.AI database of national and regional AI policies is a live repository of over 1000 AI policy initiatives, norms and rules covering 70 jurisdictions.</p> <p>POSSIBLE ACTION: In partnership with the UN and relevant UN resources, the Catalogue could be expanded and serve more widely as a global source to share good practices in AI, including standards, methods and metrics.</p>

11. Standard setting	Develop global consensus on standards for AI use across stakeholder groups by working with national standards development organizations (SDOs) - updated regularly.	Governing	12-24 months	<p>Existing standards development organisations (SDOs) should continue to play a key role as part of a multistakeholder, industry-led standards development ecosystem that remains open, transparent and consensus-based.</p> <p>The OECD.AI Expert Group on AI Risk & Accountability has been working on enabling interoperability between burgeoning AI frameworks and standards. As part of this work, the group is developing a Responsible Business Conduct (RBC) guidance for AI. RBC principles and standards set out the expectation that all businesses – regardless of their legal status, size, ownership structure or sector – identify, prevent, mitigate and account for how they address adverse impacts they cause or are linked to through a business relationship.</p> <p>The OECD Public Governance Committee is developing a <i>Framework for Responsible and Trustworthy use of AI in the Public Sector</i> to strengthen the capacities to govern with AI and help governments establish a sound governance of AI in the public sector.</p> <p>POSSIBLE ACTION: Increased cooperation with the UN (notably the OHCHR) would be welcome to facilitate global consensus and interoperability for cooperation and coordination between domestic and international state and non-state actors in the development of global AI standards.</p>
12. Norm elaboration	Convene stakeholders to assess the necessity of and negotiate non-binding and binding frameworks, treaties, or other regimes for AI.	Governing	24-36 months	<p>A number of international frameworks with implication for AI governance already exist or are in the final stages of development, such as:</p> <ul style="list-style-type: none"> • OECD Guidelines for Multinational Enterprises on Responsible Business Conduct and the UN Guiding Principles on Business and Human Rights • OECD Good Governance of Critical Infrastructure Resilience • The EU AI Act (which entails enforcement capabilities) • The Council of Europe Convention on AI (which is binding upon signatories, and open to all Members and non-Members of the Council of Europe) <p>A key goal in this context would be to seek to align international instruments as much as possible.</p> <p>POSSIBLE ACTION: The OECD could help ensure that relevant stakeholders are included in conversations about the development of frameworks for AI, including through its network of AI experts.</p>
13. Enforcement	Develop mutual reassurance schemes, information sharing mechanisms that respect commercial and national security information, dispute resolution mechanisms, and	Governing	> 36 months	<p>Responsible Business Conduct: The abovementioned OECD Guidelines for Multinational Enterprises on Responsible Business Conduct (RBC) – established in 50 countries – provide a non-judiciary dispute resolution mechanism through a network of national contact points on RBC (NCPs). The Guidelines specifically cover adverse impacts related to the development, financing, sale, licensing, trade, and use of technology; gathering and using data; and scientific research and innovation. Efforts can be made to build capacity of NCPs to better understand and resolve cases related to AI.</p> <p>AI Incidents Monitor (AIM): Through the AI Incidents Monitor (AIM), the OECD documents AI incidents and hazards to help policymakers, AI practitioners, and all stakeholders worldwide gain valuable insights into the risks and harms of AI. Over time, AIM will help to show patterns and establish a collective understanding of AI incidents and their multifaceted nature. In tandem, the OECD.AI Expert Group on AI Incidents is developing a common framework for incident reporting to enable global information sharing in a consistent and interoperable manner.</p> <p>POSSIBLE ACTION:</p>

	liability schemes/regimes.			<ul style="list-style-type: none"> • AI Incidents Monitor (AIM): The OECD could partner with the UN on a potential Version 2.0 of the AI Incidents Monitor (AIM). • Alternative Dispute Resolution (ADR): Additionally, considering solutions to help resolve cross-jurisdictional disputes in the area of AI, particularly for “small claims”, where resource to courts or human rights redress mechanisms may not be realistic, would be important and may require developing a certification scheme for ADR providers.
14. Stabilization and response	Develop and collectively maintain an emergency response capacity, off-switches and other stabilization measures.	Governing	> 36 months	The OECD Expert Group on AI Futures and associated workstream are exploring potential future AI milestones, benefits and risks, as well as potential solutions and policy approaches that can help to seize benefits while mitigating risks. With regards to risks, some of the risks being reviewed by the Expert Group could be assisted by, and may demand, stabilisation and response efforts. Some of the solutions being explored also touch on stabilisation and response. Over time, the OECD could help formalise specific recommendations and frameworks for stabilisation and response mechanisms.
15. Monitoring and verification	Elaborate oversight and verification schemes where appropriate to ensure that the design, deployment and use of AI systems is in compliance with applicable international law.	Governing	> 36 months	<p>One of the levers to assess compliance with applicable international law involves standards and certification.</p> <p>There are numerous initiatives for AI risk management frameworks in addition to existing certifications, standards and initiatives in many sectors. In order to improve the quality, comparability and interoperability of certification standards and initiatives, the OECD has developed an alignment assessment process to evaluate the alignment of initiatives with the recommendations of the OECD Due Diligence Guidance for Responsible Business Conduct, that draws from and builds on sector-specific guidance but can be applied to all businesses in all sectors of the economy, including all companies in the AI value chain.</p> <p>To note: The model proposed in Figure 1 of the interim report seems to be missing ‘Evaluation, verification and validation’ as key aspects of AI governance.</p>

OECD.AI inputs on selected “key questions for further discussion in the next phase of work” (from Interim Report, page 25)

Opportunities and enablers of AI

Question 1: Can we make AI development more inclusive by facilitating model-building ecosystems, for example through data protection and exchange frameworks, with shared access to compute?

OECD input

The OECD has established *the Expert Group on AI, Privacy and Data Governance*, which will provide insights on issues such as:

- The commonalities and possible differences in existing international AI and privacy guidelines, and whether there is a need to further facilitate interoperability;
- The meaning of key terms like "fairness", "explainability" and "transparency" for AI and data protection and privacy communities;
- The opportunities and risks around protecting privacy in an AI business models; and
- Stocktaking of various regulatory and policy approaches emerging across jurisdictions and how further global coordination on AI, data protection, and privacy governance might be achieved.

Key takeaways of an OECD report that provides a [Blueprint for building national compute capacity for artificial intelligence](#) and considers access among other policy issues include:

- AI is transforming economies and promising new opportunities for productivity, growth, and resilience. Countries are responding with national AI strategies to capitalise on these transformations.
- AI requires a special infrastructure called “AI compute”. The compute needed to train and use AI systems has grown dramatically over the last decade and continues to do so, especially for deep learning and neural networks.
- However, no country today has sufficient data on, or a targeted plan for, national AI compute capacity. This policy blind spot could jeopardise economic goals.
- This report provides the first blueprint for policy makers to assess and plan for national AI compute capacity to enable productivity gains and capture AI’s full economic potential.
- The blueprint shows how to develop a national AI compute plan along three dimensions: capacity (availability and use), effectiveness (people, policy, innovation, access), and resilience (security, sovereignty, sustainability).
- The report defines AI compute and takes stock of the indicators, datasets, and proxies countries will need to measure national AI compute capacity. It also identifies challenges to measuring and benchmarking national AI compute capacity across countries.

In addition, the OECD has considered inclusivity of AI development in its recent note [Artificially Inequitable? AI and closing the gender gap](#).

Forthcoming OECD AI Country Reviews, including that of Germany, include recommendations on widening access to data by developing data-sharing protocols and building and scaling a globally competitive computing infrastructure, which is inclusive and accessible to all stakeholders, a portion should be made accessible for use by AI start-ups and SMEs.

Question 2: Would common standards for data labelling and testing encourage AI startups to test and deploy across more countries and regions?

OECD input

The OECD's report on [Regulatory sandboxes in artificial intelligence](#) notably covers the potential of AI standards as policy tools and sources of innovation and competition. The OECD report [Artificial Intelligence and international trade](#) encourages considering and prioritising the adoption of international standards to enable easier use of AI across countries and notably analysis issues affecting and instruments facilitating cross-border data transfer.

In addition, the OECD is currently exploring how AI can support its own work on environmental data collection, indicators and measurement with a view to drawing lessons learned for policy makers. Upcoming work will focus on the role AI can play in enhancing monitoring and compliance with existing regulations and accelerating permitting processes; all with important implications for environmental policy design.

Question 3: What mechanisms would promote equitable access to compute and privacy-preserving sharing of datasets across stakeholders and member states?

OECD input

Mechanisms that promote equitable access to computing and the exchange of datasets that preserve privacy are critical through data governance frameworks that manage shared data and promote collaboration through common rules on security, privacy and confidentiality. The OECD has recently conducted analysis on [Emerging privacy-enhancing technologies \(PETs\)](#), which allow information to be collected, processed, analysed, and shared while protecting data confidentiality and privacy.

The OECD is conducting an analysis on the 'legal bases for the secondary uses of health data' – that includes the use of data for training AI solutions. The outcomes of this work will highlight the needs for (1) technologies (such as PETs) and (2) harmonised policies across sectors and member states that enable access while ensuring protections are in place – as referenced in the [OECD Recommendation on Health Data Governance](#) (2016).

Question 4. How can we grow and spread AI talent? Can UN entities or other institutions facilitate exchange of students, joint PhD programmes, and cross-domain (health and AI, agriculture and AI) talent development?

OECD input

The OECD has analysed national programmes to grow and spread AI talent as documented in the [OECD programme on Work, Innovation, Productivity and Skills](#).

Public administrations can undertake institutional learning to ensure that they equip their employees with the knowledge and skills required to effectively govern and use AI technologies as part of delivering more effective government services. OECD's work on digital skills and talent for the public sector provides a framework to support countries in creating the right working environment, securing the right skills, and evolving the right workforce. The OECD can assist in developing a shared curriculum for public servants to acquire core skills in using AI responsibly, based on the OECD Digital Talent and Skills Framework.

Question 5. How can international collaboration harness AI talent, data and compute for scientific research and for the SDGs?

OECD input

Among other, international collaboration allows us to bring together experts from different countries, foster the exchange of knowledge and skills that can lead to faster and more effective innovation, including by facilitating access to large data sets needed to train AI models.

The OECD has gathered extensive analysis of the potential of [Artificial intelligence in science](#), which highlights – among others – the importance of ‘knowledge bases’ for AI in science, the necessity for public investment in research and development in AI.

The OECD work on Governing with AI and its focus across various policy areas will provide further knowledge on talent and data available within the public sector to address the SDG’s.

Early OECD work on the impact of AI on the environment includes a publication on [Measuring the environmental impacts of artificial intelligence compute and applications](#) detailing existing measurement tools and challenges governments face in quantifying this impact. An ongoing review of the 2010 [OECD Recommendation on Information and Communication Technologies and the Environment](#) takes into account the impact of AI on the environment and policy tools to mitigate it, as well as the potential of AI solutions to promote sustainability. These issues were the focus of the recent Green Growth and Sustainable Development (GGSD) Forum on “Navigating the twin transition: going green and digital” (21-22 November 2023, Paris), which drew on relevant workstreams under the OECD’s Horizontal Project “[Net Zero+: Building Climate and Economic Resilience in a Changing World](#)” (Net Zero+).

Question 6. How can we incentivize governments and the private sector to invest in other core infrastructures that drive AI development around the world?

OECD input

One of the key infrastructures needed to drive AI development around the world is the computing infrastructure, which is dependent on the availability of semiconductor chips – [government support in the semiconductor value chain](#) as well as the [Vulnerabilities of semiconductor supply chains](#), have been extensively analysed by the OECD.

Increasing computing infrastructure could lead to an increase in greenhouse gas (GHG) emissions due to the consumption of (often fossil-fuel-based) electricity by data centres. This is an environmental risk of AI that needs to be well understood and managed.

Early OECD work on the impact of AI on the environment includes:

- A publication on ‘[Measuring the environmental impacts of artificial intelligence compute and applications](#)’, describes existing measurement tools and challenges governments face in quantifying this impact.
- An ongoing review of the 2010 [OECD Recommendation on Information and Communication Technologies and the Environment](#), which takes into account the impact of AI on the environment and policy tools to mitigate it, as well as the potential of AI solutions to promote sustainability.

The recent Green Growth and Sustainable Development (GGSD) Forum on “Navigating the twin transition: going green and digital” (21-22 November 2023, Paris), which drew on relevant workstreams under the OECD’s Horizontal Project “[Net Zero+: Building Climate and Economic Resilience in a Changing World](#)” (Net Zero+). Further OECD analysis deals with other core infrastructures for AI, including

[telecommunications](#) and connectivity infrastructures ([broadband](#), fibre and satellite).

Risks and challenges of AI

Question 7. What is the best path to reaching consensus on identifying, classifying, and addressing AI risks?

OECD input

OECD consensus on [classifying AI risks](#)

AI enables new ways to learn, work, play, interact and live. In many cases, AI can make the lives of people easier (autocorrect), more secure (facial recognition on smartphones) or even more exciting (online gaming). However, as AI spreads across sectors, different types of AI systems deliver different benefits, such as financial prosperity, or benefits for tackling climate change (such as greater reliability of modelling and projections, risk assessment, or resource use) and risks in areas such as harmful bias, human rights including safety, security, and privacy, labour and intellectual property rights human rights, , or to the environment.

Both the risks and benefits bring policy and regulatory challenges. While most people would agree that governments should take steps to protect citizens, they probably also agree that progress should not be impeded by excessive regulation. Nor do they want to relinquish individual rights.

To add another layer of complexity, all of these considerations do not simply depend upon the technology itself. Factors such as context, and which stakeholders are affected by an AI system's tasks and outputs, are also part of the equation.

Consider the differences between a virtual assistant, a self-driving vehicle and an algorithm that recommends videos for children. It is easy to see that the benefits and risks can be measured not only by what the technology can do but also by the context in which it acts and the stakeholders who are affected by the outcomes. This complexity is even more obvious in the example of facial recognition technology, which is great for smartphone security but can threaten human rights and liberties in other circumstances.

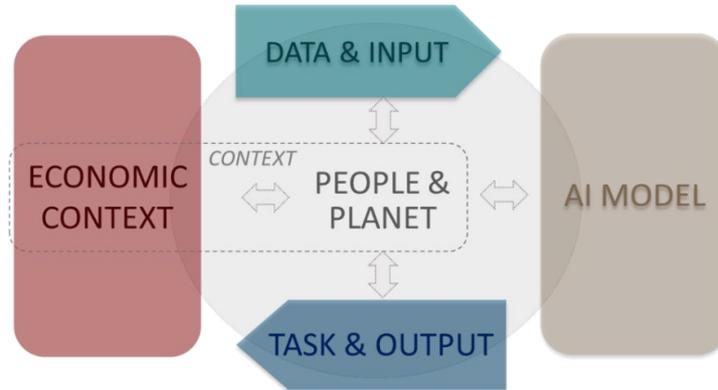
The OECD Framework for Classifying AI Systems

[In February 2022, the OECD released a user-friendly framework](#) to guide policy makers, regulators, legislators and others as they characterise AI systems for specific projects and contexts. The framework links AI system characteristics with [the OECD AI Principles](#), the first set of AI standards that governments pledged to incorporate into policy-making and that promote the innovative and trustworthy use of AI.

The framework's key dimensions structure AI system characteristics and interactions

The dimensions help users to sort out critical details about an AI system, how it works, where responsibilities lie, and how its outcomes may affect people and the environments in which it operates. The framework classifies AI systems along the following dimensions: People & Planet, Economic Context, Data & Input, AI Model and Task & Output. Each one has its own properties and attributes that help assess policy considerations of particular AI systems. Stakeholders are involved in or affected by AI systems, while AI actors play active roles according to each dimension and throughout an AI system's lifecycle.

Figure 1. Key high-level dimensions of the OECD Framework for the Classification of AI Systems



Question 8: What kind of mechanism could best support industry participation in international governance of AI? Which of the normative, policy and information instruments that exist today could support coherence in technology governance across governments, private sector and civil society?

OECD input

The OECD is developing [Due Diligence Guidance for Responsible Business Conduct in AI](#), in partnership with industry, academia, trade unions and civil society.

The OECD is developing a Framework for Anticipatory Governance of Emerging Technologies, which aims to i) embed values in the innovation process ii) assess technology with strategic intelligence, iii) engage stakeholders and society iv) build agile regulation, and v) cooperate internationally. At the end of the current consultation process, the Framework is set to be adopted by a meeting of the OECD Committee on Scientific and Technological Policy in April 2024. Preliminary considerations for the framework are publicly available [here](#).

OECD’s work on [Innovative Citizen Participation](#) explores evolving deliberative, collaborative, and participatory decision-making within governments to effectively engage citizens and stakeholders in sourcing ideas, co-creating solutions, and addressing complex policy challenges, such as the governance of AI. Lessons and best practices can be extended and further explored to national, regional, and international governance of AI.

International governance of AI

Since 2016, the OECD has been an integral part of international governance of AI, adopting the first intergovernmental standard on AI. The OECD AI Principles were adopted by OECD member countries in 2019 and also constitute the basis for the G20 AI Principles. Its revised definition of an AI system adopted by the OECD in December 2023 has already been used in the EU AI Act and the Council of Europe’s treaty on AI, as well as referenced in this Interim Report. The OECD stands ready to continue to collaborate with the United Nations and other international organisations to jointly shape international governance on AI.