Organisation for Economic Co-operation and Development

DSTI/CDEP(2022)14/FINAL

Unclassified

English - Or. English

15 November 2022

DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INNOVATION
COMMITTEE ON DIGITAL ECONOMY POLICY

Harnessing the power of AI and emerging technologies

Background paper for the CDEP Ministerial meeting

JT03507755

Foreword

This paper explores the opportunities and risks posed by AI and emerging technologies, including risks to human rights, fairness and human agency. It builds on OECD's extensive work on AI, data governance and connectivity to support policy makers in the process of developing forward-looking policies and adapting governance frameworks to keep pace with these technological developments and ensure they are trustworthy.

The paper provides background to support the discussions on Theme 4: *Harnessing the Power of AI and Emerging Technologies* of the Ministerial meeting of the Committee on Digital Economy Policy, taking place on 14-15 December 2022 in Gran Canaria, Spain. It informs the sessions on "The OECD AI Principles – impact on the global policy landscape" and "The future of simulated environments and immersive technologies" of the Ministerial meeting.

This paper was written by Karine Perset, Christian Reimsbach-Kounatze, Inmaculada Cava Ferreruela, Kulani Abendroth-Dias and Luis Aranda, under the supervision of Audrey Plonk, Head of the OECD Digital Economy Policy Division. It benefitted from the inputs of Hanna-Mari Kilpelainen, Verena Weber, Elizabeth Thomas-Raynaud, Gallia Daor, Adam Mollerup and colleagues from the OECD Employment, Labour and Social Affairs Directorate (ELS), the Education Directorate (EDU), the Directorate for Financial and Enterprise Affairs (DAF), the Centre for Entrepreneurship, SMEs, Regions and Cities (CFE) and the Directorate for Public Governance (GOV). The report also benefitted from the inputs of delegates for the OECD Committee on Digital Economy Policy, including the Civil Society Information Society Advisory Council (CSISAC) and Business at the OECD (BIAC). Shellie Phillips, Angela Gosmann, Sebastian Ordelheide and Misha Pinkhasov provided editorial support. The Ministerial meeting and related work were generously supported by the Government of Spain.

This paper was approved and declassified by written procedure by the Committee on Digital Economy Policy on 26 October 2022 and prepared for publication by the OECD Secretariat.

Note to Delegations:

This document is also available on iLibrary as:

OECD (2022), "Harnessing the power of AI and emerging technologies: Background paper for the CDEP Ministerial meeting", *OECD Digital Economy Papers,* No. 340, OECD Publishing, Paris, https://doi.org/10.1787/f94df8ec-en.

Table of contents

FIGURES

Executive summary

In 1950, Alan Turing first asked whether machines can think. Since then, innovation in computational capabilities, connectivity and data availability have led to breakthroughs in which machines learn from large amounts of data to generate predictions and other types of outputs.

People now live alongside digital technologies in the physical and virtual worlds. These include machines pre-programmed to follow a precise set of rules, or that are fully autonomous and can operate without human intervention. They include immersive environments that combine features of the physical and virtual worlds to create realistic experiences, such as surgical training, that would be difficult to reproduce in real-world settings. Behind these and other innovations in the early phases of implementation are complex mathematical models trained on large computers with vast amounts of data to emulate human-like cognitive functions, i.e. Artificial Intelligence (AI).

AI and emerging technologies are reshaping societies with opportunities, but also risks. Most countries identify these technologies as priorities and levers for economic growth, while noting the need to adapt governance frameworks. There are many types of AI systems, from voice recognition, to chatbots, product and content recommendations and driverless cars, many of which also raise risks. These include bias and discrimination, the polarisation of opinions at scale, upheaval in job markets due to the automation of tasks, privacy infringement, widespread surveillance, threats to security and safety and increasing inequalities from concentration of power. Immersive environments also pose risks, from online addiction to virtual assault, to the challenge of assigning liabilities in a world of avatars.

Public policy frameworks must adapt to meet the governance imperatives of AI and emerging technologies and protect established rights. Policies in these areas are key to building a strong digital ecosystem in which governments, businesses and individuals can reap the benefits.

Harnessing the power of AI and emerging technologies: Background paper for the CDEP Ministerial meeting

AI and emerging technologies offer tremendous opportunities for well-being, productivity, growth and solving pressing societal challenges. However, they also pose risks to human rights, fairness and human agency, among others. Many countries recognise the need to develop forward-looking policies and adapt governance frameworks to keep pace with these developments and to leverage technological benefits while mitigating risks. This paper builds on OECD's extensive work on AI, data governance and connectivity to support policy makers in this process. It highlights the importance of co-operating internationally to ensure that emerging technologies are trustworthy and calls for building a common understanding of AI and emerging technologies, sharing good practices and creating the evidence base to inform policy design, implementation and evaluation.

The promise and peril of AI and emerging technologies

AI permeates economies and societies

Products and services powered by AI are already a crucial part of most people's routines, whether they realise it or not (Box 1). Consider AI systems that recommend articles, friends and products; recognise faces, speech and objects; and power chatbots or optimise traffic routes. Other examples include AI supporting climate action, disaster prevention and public service delivery, with data gathered by Internet-of-Things (IoT) devices and connected sensors (OECD, 2020[1]). Predictive models can anticipate service demand by combining historical and context-sensitive data from multiple sources (e.g. weather, mobile, big data) to allow for a strategic deployment of vehicles or other productive assets (OECD, 2019[2]).

Box 1. AI in road safety

Road-safety is one example of AI's potential to solve complex challenges when it is developed and used in a trustworthy way. Each year, about 1.35 million people are killed in car crashes and 50 million are seriously injured (World Health Organisation, 2018[3]). Speeding, distractions, impaired driving and not wearing a seatbelt are the top causes of automotive deaths. High-visibility traffic enforcement – such as officers in marked cars – can improve behaviour on the road. Drivers often put away their phones, put on their seatbelts and drive slower in their presence (Mohn, 2022[4]). But it is not possible for officers to patrol every road all the time.

Technological advancements in AI, including machine learning, are being harnessed to enhance road safety around the world. For example, intelligent speed-assistance technology, which uses AI to manage a car's speed via in-vehicle cameras and maps, became mandatory in all new vehicles in the EU as of July 2022.

In Barcelona, Spain, a trial used computer vision technology on city buses to map hotspots where accident risks were highest (International Transport Forum, 2021[5]). AI scans raw video feeds for road layouts, positions of pedestrians and vehicles, streetlights and speeds. Where precise, long-term and relevant data exist, AI systems identify dangerous locations to reduce the risk of future accidents. In Australia, companies are coupling high-resolution imaging with machine learning to identify dangerous driving behaviours. In addition to monitoring vehicle speed, algorithms can determine with high probability whether a particular driver is engaged in risky behaviour; for example, assessing if the driver is looking down to text someone or is holding something besides the wheel. In the first two years of the pilot, fatalities decreased 22% and handheld phone use while driving decreased 80% (Mohn, 2022[4]).

Looking forward, the potential of self-driving vehicles to enhance road safety is being studied and tested in several countries. In the US, 94% of traffic crashes were attributed to human error (National Highway Traffic Safety Administration, 2015[6]). By reducing the risk of human error, self-driving vehicles could significantly reduce the risk of automotive crashes.

In manufacturing, AI is being leveraged to create digital twins, i.e. digital representations of a physical objects or systems, which simulate different scenarios and predict outcomes to inform decisions. Digital twins can optimise the quality of manual production tasks, provide real-time training and perform predictive maintenance and repairs with minimum downtime. AI

However, the benefits of AI-enabled technologies are accompanied by real risks. For example, what will happen to truck, ride-hailing and delivery drivers if we switch to self-driving vehicles at scale? What are the implications of facial-recognition systems that detect the attention and emotion of drivers? How can we safeguard privacy when vehicles with self-driving capabilities generate and transmit floods of data?

How can we avoid discrimination against protected groups in services like car insurance and ensure transparency and explainability? How can we guarantee the robustness, security and safety of vehicles with self-driving capabilities?

AI brings opportunities and risks in all sectors, including those with traditionally high women participation such as education, healthcare and customer service.[1] Unchecked AI can create dangers and misalignment with human and democratic values: discrimination through the automation or amplification of biases; polarisation of opinions; upheaval in job markets with the automation of tasks; privacy infringement; widespread surveillance; and the concentration of power. We must leverage AI opportunities and address AI risks in a timely manner to ensure AI is beneficial for people and planet. In addition, because AI is global and impacts everyone and because AI developments outpace policy, the development and use of AI call for a coordinated policy response based on international, multi-disciplinary and multi-stakeholder co-operation.

Virtual worlds offer real benefits and costs

As a general-purpose technology, AI enables and supports other emerging technologies, such as immersive virtual environments, referred to as "virtual worlds" and the "metaverse", a concept popularised by Neal Stephenson's science fiction novel *Snow Crash* (Stephenson, 1992[7]). Today, these immersive environments are based on augmented reality (AR), virtual reality (VR), mixed reality (MR) and other extended reality (XR) technologies that enhance the realism of virtual experiences, blurring the lines between the physical and digital worlds. They also increasingly rely on AI-enabled prediction and personalisation, interaction support, speech recognition and language translation and low-latency connectivity to augment the immersive experience.

Immersive environments combine features of the physical and virtual world: they provide realistic experiences important for communication (e.g. in education) and experimentation (e.g. in medicine) with the comfort, safety and the cost- and time-saving value of not having to travel. They are viewed as promising to support advances in education, healthcare, manufacturing, banking, marketing and entertainment (Verizon, 2022[8]; Healthcare Insights, 2022[9]). In education, they can help develop dynamic, safe and remote learning environments for students and professionals and deliver realistic environments to support learning and distance education (Melchor-Couto, 2019[10]). These virtual worlds can be consumer-oriented, entertainment-oriented, corporate, industrial, or private and likely to inter-connect.

However, immersive environments create challenges. By providing hyper-realistic experiences, where virtual sensations can seem real (including using tactile or haptic devices), immersive technologies could further exacerbate issues like online addiction, discrimination, cyber-bullying, assaults and abuses and inappropriate or illegal content. These concerns and broader challenges raise questions about the applicability of prior experience in accountability and user-protection from the governance of the Internet and emerging technologies in general.

In addition, immersive environments promise to generate significant economic activity as marketplaces where people use virtual avatars to buy virtual real estate, services, clothing, or artwork that have real-world value. This raises issues pertaining to intellectual property rights, taxation and money laundering. With physical reality transposed to the virtual world, developers, deployers, regulators and users must understand and address existing and emerging challenges, recalling that human motivations are being replicated in a virtual world – and with them, their biases and aggressive behaviours. These must be prevented and mitigated and actors in physical and immersive spaces should be held accountable according to their roles. Ongoing efforts to establish governance frameworks for immersive environments include the Defining and Building the Metaverse multistakeholder initiative by the World Economic Forum (WEF) and efforts by the ("eXtended Reality") XR Association and the Metaverse Standards Forum.

Box 2 provides an overview of policy research and legal instruments developed by the OECD Committee on Digital Economy Policy to support the governance of AI, data and connectivity that can inform the governance of emerging digital and immersive technologies.

Box 2. Overview of key policy research and legal instruments developed by the OECD Committee on Digital Economy Policy

Policy research

OECD (2021-2022), *Horizontal Project on Data Governance for Growth and Well-Being.*

OECD (2022[11]), *Broadband networks of the future*, *OECD Digital Economy Papers*, No. 327.

OECD (2022[12]), *OECD Framework for the Classification of AI systems*, *OECD Digital Economy Papers*, No. 323.

OECD (2019[13]), *Artificial Intelligence in Society.*

Legal instruments

OECD (2021[14]), *Recommendation on Enhancing Access to and Sharing of Data.*

OECD (2021[15]), *Recommendation of the Council on Broadband Connectivity*.

OECD (2019[16]), *Recommendation on Artificial Intelligence.*

OECD (2013[17]), *Recommendation concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*

Enablers of AI and emerging technologies

Over the past few years, the availability of large amounts of data, breakthroughs in machine learning, high-quality connectivity and the expansion of computation power dramatically increased the capability, availability, growth and impact of AI and emerging technologies. Policies in these areas are key to building an enabling ecosystem where governments and citizens can enjoy their benefits.

Data access, sharing and governance

Access to and sharing of data are critical to enable AI's benefits across sectors. Policies must encourage data access and sharing while addressing associated risks, for countries to harness the full potential. There are concerns about the concentration of technology and financial resources in the hands of few companies and nations, reflected in the control of large data sets and computational power to process them. This has follow-on costs to societies in "winner-takes-most" markets, including slowing productivity and growing gaps in AI adoption between large companies and SMEs and between large companies and the public sector. At the same time, there are tangible reasons for companies' reluctance to share data. In addition, some large firms proactively contribute to large public datasets and to improving AI data quality for public use.

In many countries, more efforts are needed to reduce siloes in data governance, despite growing efforts to govern data across policy domains, like competition and trade and across sectors, like science, healthcare and public administration. As data fuel AI across a range of economic and social activities in the private and public sectors, data governance is becoming pervasive across policy domains and requires coordinated, whole-of-government approaches, while remaining sensitive to specific contexts.

Yet in many countries, policymakers and regulators face difficulties finding common definitions and common ground in discussions, co-operation and coordination on data governance, at national and international levels. They focus on aspects relevant to their policy domains and jurisdiction. Agreeing upon common approaches and aligning efforts across borders will play a role in the responsible use of data and AI across sectors and borders (OECD, forthcoming[18])

Therefore, data should be governed to maximise its benefits while addressing risks and challenges, including protecting the rights of individuals and organisations. This requires comprehensive policy to address cross-cutting challenges, while accounting for the specificities of data governance in domains like trade or competition (OECD, forthcoming[19]). These include:

- **Balancing the trade-offs between data openness and control.** The more openly data is accessed, shared and re-used (for example, with open data), the higher its potential social and economic benefits, but also the greater the associated risks.
- **Addressing potentially conflicting interests and regulations.** Data collected and used to inform AI systems are often (co-)created by the interaction of many stakeholders in the global data ecosystem, in some cases without them being aware. Facilitating data access and sharing for AI requires disentangling and reconciling these interests and data-governance frameworks.
- **Aligning incentives for investment in data and its re-use.** While the marginal costs of transmitting, copying and processing data can be close to zero, substantial investment is often required to generate and collect data and enable data sharing and re-use for AI. Fair distribution of the benefits from data can help address incentive challenges.

Machine learning models and techniques

AI and emerging technologies use machine learning models and techniques to learn in an automated manner through patterns and inferences rather than explicit instructions from a human. Machine learning approaches teach machines to reach an outcome by showing them many examples of correct outcomes. However, they can also define a set of rules and let the system learn by trial and error. Machine learning contains techniques that have been used by economists, researchers and technologists for decades (ranging from linear and logistic regressions to decision trees and principal component analysis), but also more recent technological developments including deep neural networks (OECD, 2019[13]).

Machine learning models and techniques are key to the expansion of AI and emerging technologies. They have led to significant progress in research areas such as natural language processing, computer vision and robotics.

Connectivity extension and enhancement

Connectivity allows the transfer of large volumes of data in real or quasi-real time, while computing infrastructure (hardware and software) executes the mathematical operations needed to calibrate or "train" an AI system and infer its results (OECD, forthcoming[20]). The combination of high-quality connectivity, data, computing infrastructure and AI technologies continues to enable innovative and disruptive new services.

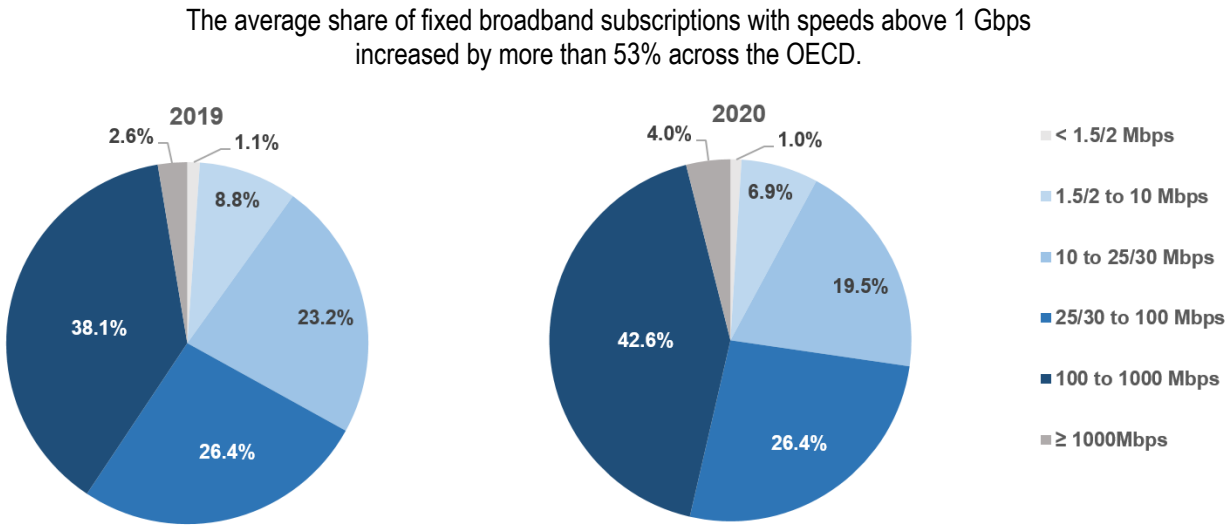
High-quality connectivity is characterised by features such as high speed and responsiveness (i.e. low latency, or delay introduced by the network) and reliability. High-speed connectivity enables richer and more interactive applications with ultra-high-definition video, holograms and augmented and virtual reality applications. People are expected to soon be able to gather for live events such as sports and concerts virtually, with 360-degree views from anywhere they wish.

In addition, responsive and reliable connectivity play a role in enabling tactile digital environments, where people interact virtually with each other and the environment. This is critical to allow autonomous or semi-

autonomous systems like vehicles, drones and robots to respond immediately to changes in their environment. Responsive and reliable connectivity supports extremely delay-sensitive applications such as tele-surgery that, combined with AI image recognition, can maximise the efficiency of medical resources and allow the best medical services to be provided in rural areas.

This enhanced connectivity is made possible by next-generation mobile networks like 5G (and eventually 6G) and fixed networks like fibre, capable of speeds of gigabits per second, latencies as low as 1 millisecond and "hyper-connectivity" by a massive myriad of devices. Evidence shows that operators are increasingly investing in these high-quality connectivity technologies and AI to manage cutting-edge networks more efficiently and offer end-user services. OECD broadband data (Figure 1) shows that the share of gigabit fixed broadband offers (i.e. broadband with speeds above 1 gigabit per second) across the OECD grew by more than half, from 2.6% at the end of 2019 to 4% by December 2020 (OECD, 2022[11])

Figure 1. Fixed broadband subscriptions by speed tiers, 2019-2020



Note: Simple average of countries where data was available (i.e. 35 out of 38 OECD countries).
Source: (OECD, 2022[21]).

Despite these developments, high-quality networks require continued investment to face increasing demand and provide ubiquitous coverage. In Europe, the funding gap to meet the 2025 Digital Agenda and Gigabit Strategy objectives is projected at EUR 384 billion (USD 453 billion), of which 66% would require public policy intervention (European Investment Bank, 2018[22]; OECD, 2021[23]).

The 2021 OECD Recommendation on Broadband Connectivity (OECD, 2021[15]) is a policy and regulatory roadmap that puts forward principles to extend connectivity and enhance the quality of broadband networks. It complements G20 Guidelines for Financing and Fostering High-Quality Broadband Connectivity for a Digital World (G20/OECD, 2021[24]), developed with OECD support. These instruments highlight the importance of policies that favour the deployment of resilient networks that can ensure high-quality communication services in cases of power outages, natural disasters, or other disruptions.

Expansion of computational power

Computation is critical to the development of AI and emerging technologies. Policymakers have begun to realise the strategic importance of specialised AI compute. Since 2010, the prominence of a type of AI known as deep neural network dramatically increased the size of machine learning systems and consequently their compute demands. Satisfying this demand was partially enabled by transitioning from

general-purpose processors, such as Central Processing Units (CPUs), to specialised processors that support more efficient compute execution, i.e. require less energy and time per computation. Today, machine learning systems predominantly train on specialised processors optimised for operations commonly used in, but not limited to, machine learning. These processors include Graphics Processing Units (GPUs), Tensor Processing Units (TPUs) and Neural Processing Units (NPUs). Executing such AI workloads (e.g. training machine learning systems) on general-purpose hardware is less efficient.

Access to AI compute is necessary to advance and diffuse AI. Ensuring countries have sufficient AI compute to meet their needs is critical to capturing AI's full economic potential. Yet the hardware, software and related compute infrastructure that make AI advancements possible receive less attention in policy circles than other enablers, like data. Policymakers have begun to pay attention to the risk of an "AI compute divide" growing both within countries – for example between the private sector and academia (Ahmed and Wahed, 2020[25]) – and between countries – for example between developing and developed nations. A compute divide risks creating gaps between those who have the resources to develop and train the large-scale AI models that lead to competitive advantage, inclusive growth and productivity gains in a global digital economy and those who do not. It requires a measurement framework for national AI compute.

Investing in infrastructure and hardware purpose-built for AI is challenging given complex supply chains, as illustrated by recent shortages in the semiconductor industry (Khan, Mann and Peterson, 2021[26]). Semiconductors, also known as integrated circuits or computer chips are the "brains of modern electronic equipment, storing information and performing the logic operations that enable devices such as smartphones, computers and servers to operate", which are critical for the development and use of AI systems (OECD, 2019[27]). In any electronic device, there can be multiple semiconductors fulfilling specific functions, such as CPUs or chips specifically designed for power management, memory, graphics and more.

Demands on semiconductor supply chains have grown, especially as digital and AI-enabled technologies have become more commonplace, including in Internet-of-Things (IoT) devices, smart energy grids and self-driving and electric vehicles. The semiconductor supply chain is also concentrated, with the top five companies representing half of global revenue in 2018 (OECD, 2019[27]), making it vulnerable to supply-side market shocks. Leaders have associated semi-conductor supply shortages to reduced activity in some industries and resulting cost and inflationary pressures (OECD, 2021[28]).

The growing computational needs of AI systems are also raising sustainability concerns. At the same time, higher computational power could result in efficiency gains and reduced energy consumption in some sectors. The environmental impacts of AI compute and applications should be further measured and understood (OECD, 2022[29]).

Building trustworthy AI and emerging technologies

Rapid developments in AI make it difficult for policymakers, regulators and other governance bodies to keep pace and implement forward-looking policies and governance frameworks that enhance potential and mitigate risks. General-purpose technologies of global application like AI cannot be realised and governed by a single country or economic actor. The computational infrastructure, skills and data needed to develop and deploy a single AI system often cross jurisdictional borders. Therefore, international, multi-disciplinary and multi-stakeholder co-operation and coordination are needed to ensure that AI innovations improve the well-being of people in areas such as education, public safety, health and work-life balance. Restrictive policies affecting cross-border data, skills and technology flows might limit the development and uptake of AI systems. Taking a balanced approach to promoting AI and the benefits of emerging technology for innovation and well-being gains while managing risks is key. The opportunity costs of not using technologies like AI where they can provide benefits and insights must also be considered. In this regard,

risk-based approaches are being proposed by players like the US National Institute of Standards and Technology (NIST) and regulatory initiatives like the EU AI Act.

Shaping a common understanding of trustworthy AI and emerging technologies

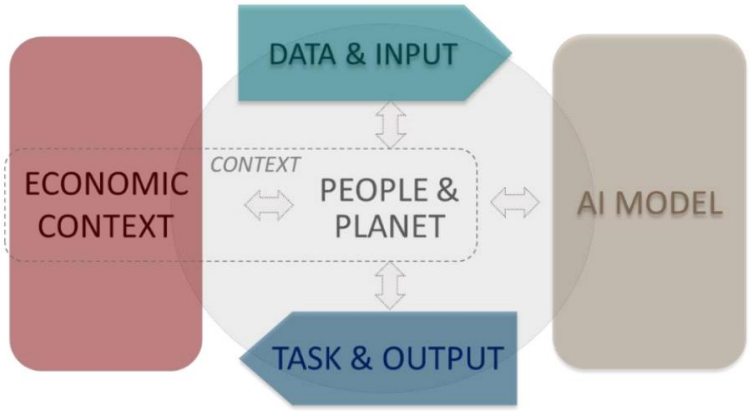
The OECD AI Principles provide a common denominator for AI governance (OECD, 2019[16]). They define AI as:

> *a machine-based system that is capable of influencing the environment by producing an output (predictions, recommendations or decisions) for a given set of objectives. It uses machine and/or human-based data and inputs to (i) perceive real and/or virtual environments; (ii) abstract these perceptions into models through analysis in an automated manner (e.g. with machine learning), or manually; and (iii) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy (OECD, 2019[30]).*

By adopting the OECD AI Principles, countries agreed to a common set of priorities to promote and harness the power of trustworthy AI, namely that AI systems: (1) benefit people and the planet; (2) respect democratic values and human rights, including privacy and fairness; (3) be transparent and explainable; (4) be robust, secure and safe; and (5) hold AI actors accountable for their proper functioning.

Subscribing countries also agreed to use risk-based approaches to govern AI, with policies and standards that treat different AI applications differently depending on the risks they pose. The OECD's classification framework (Figure 2) allows users to focus on the specific risks of an AI application based on its technical characteristics (e.g. risks to safety and privacy in the case of autonomous vehicles). This user-friendly framework – based on an AI system's characteristics and policy implications – facilitates a more nuanced and precise policy debate.

Figure 2. OECD Framework for the Classification of AI Systems



Source: (OECD, 2022[12]).

The framework can help develop policies and regulations by providing a baseline to advance understanding of AI based on system characteristics. This, in turn, informs registries or inventories of AI systems; nuanced, sector-specific approaches to implement the AI Principles, such as in transportation, healthcare, or finance; and the development of risk assessments and incident reporting mechanisms that are interoperable across jurisdictions. As AI knows no borders and many actors operate internationally, the OECD is using this classification to facilitate interoperability across countries in assessing risk alongside national and international experts and stakeholders, including standards bodies.

Building and exchanging knowledge about national AI and emerging technology strategies and initiatives

National strategies and policies that focus specifically on AI are a relatively new phenomenon that Canada began in 2017. Today, AI strategies and policies are found in over 60 countries. National strategies should build on lessons around a shared understanding of AI, building data-protection frameworks and investing in connectivity. The OECD.AI Policy Observatory's database of national AI policies and strategies from the countries can provide information and inspiration to design, implement and consult on countries' AI policies and strategies, as underlined in the sections below. The OECD also collects case studies and conducts analysis on how governments can use AI to design and deliver better policies and services, which feed into the efforts of the Policy Observatory. National AI strategies and policies can help implement the OECD AI Principles and harness the power of AI. National strategies reflect a country's vision, context and priorities in AI. The national priorities of most countries with an AI strategy can be mapped broadly to five recommendations to governments included in the OECD AI Principles. Examples of how countries are implementing these recommendations to harness the power of AI are listed below.[2]

Investment in research and development

Countries are funding national AI-related research institutes and projects through grants; consolidating AI research networks and collaborative platforms; prioritising AI investments in specific sectors; pursuing AI-related mission-oriented innovation policies; and procuring AI systems for the public sector. Budgets for AI R&D vary across countries. Since 2020, the United States dedicates USD 1 billion or more annually to non-defence AI R&D and created national AI research institutes. The EU Horizon 2020 programme committed EUR 1.5 billion to AI research over two years and expected an additional EUR 20 billion in 2020 from the private sector and member states, with the Horizon Europe programme continuing these efforts.

Digital ecosystems

As measured by the OECD Open, Useful and Re-usable data (OURdata) Index (OECD, 2020[31]), open access to public-sector data continues to be a priority as national data strategies focus on AI to foster a robust digital ecosystem. Policies to promote access to public data (including open government data) and initiatives that enable private-sector data sharing include data trusts, data dams and data spaces. Several countries have centralised, open-government data platforms, such as anonymised government health records and satellite data (e.g. Chile, Norway, Portugal, Spain and the US) (OECD, 2021[32]). Others, like the UK and the EU are looking for ways to incentivise data sharing by the private sector. Korea plans to incentivise data sharing by the private sector by estimating and pricing data value and rewarding its sharing. Most national AI strategies also recognise the importance of trustworthy AI in the public sector and the need to build upon and inform smart city strategies (OECD, 2020[1]).

Alongside data, machine learning models and connectivity, AI computing capacity has emerged as an enabler for AI and economic growth and competitiveness. National policies (e.g. EU, UK) prioritise investments in high-performance computing and cloud computing to increase AI development and use, alongside the development of domestic semiconductor manufacturing capacity.

At the same time, countries are becoming aware of the high energy consumed by large computational infrastructure. Alternatives, such as the use of clean energies and more efficient algorithms, are being explored. Conversely, the role of AI in clean and secure energy systems has been highlighted in international discussions. As noted above, connectivity is a key element of the AI ecosystem, with high-capacity, broad-coverage networking being critical to AI innovation, deployment and operation.

Enabling policy environments

AI-related legislative proposals are gaining traction in many countries, with particularly strong momentum in heavyweights like the EU, the US, the People's Republic of China and Brazil. Policymakers are developing governance and accountability frameworks such as the proposed US Algorithmic Accountability Act and the EU AI Act. Existing laws are also being re-interpreted in an AI context, like workplace rule regarding hiring, performance management, training and lay-offs, with rulings in the US and other countries (Salvi, Wyckoff and Vourc'h, 2022[33]).

The focus is not only on legislation, which requires thoughtful preparation to stand the test of time against a rapidly evolving technology like AI. Governments and other stakeholders are using various policy instruments – particularly standards. The NIST was directed by the US Congress to develop a risk management framework and issued a second draft framework for comment in August 2022 (NIST, 2022[34]). The European Committee for Standardization (CEN) and the European Committee for Electrotechnical Standardization (CENELEC) received a draft request by the European Commission in May 2022 to create AI standardisation initiatives with a view to harmonising standards by the time the proposed EU AI Act regulation is applicable (planned late 2024/early 2025).[3]

Policymakers and market actors are establishing regulatory-experimentation ecosystems to test AI in controlled environments, such as sunset provisions, innovation test beds, policy prototyping, innovation hubs, regulatory beaches and sandboxes. These initiatives aim to improve conditions for AI to strive and scale up, remove market barriers to AI adoption by businesses and enable a feedback loop for the fine-tuning of regulation. AI regulatory sandboxes are a way to generate much-needed evidence to inform public policies, with firms benefiting from a waiver from liability to test projects in a controlled environment under regulatory supervision or to co-create, prototype and test specific policies. AI regulatory sandboxes can also be established by private or multistakeholder initiatives.

Other policy initiatives include designing participatory processes and public consultations on AI strategies (e.g. Chile); connecting emerging companies with business opportunities through networking and collaborative platforms (e.g. Canada, Colombia, Germany, Slovenia); providing tailored advice to support businesses as they scale up (e.g. Germany, Finland, AI4EU); and improving companies' access to funding, including for SMEs (e.g. UK, Türkiye, Brazil) (OECD, 2021[35]). The development and uptake of trustworthy AI needs to encompass companies of all sizes and most governments recognise that SMEs might require additional support and guidance.

Human capacity and labour market transformation

Policymakers are establishing education programmes related to AI, including: developing vocational training and lifelong learning programmes in AI-related fields to help citizens keep up with technological and societal changes; providing financial and non-financial support to retrain and attract top AI talent, including migration quotas and new visa routes; fostering academic partnerships between public and private AI research institutions; using AI to match people to jobs based on skills; and monitoring the impact of AI on the labour market for policy intervention.

Regulating the use of AI in recruitment and hiring, which could reinforce or augment existing biases, is being considered in some jurisdictions, such as in the proposed EU AI Act.

International co-operation

Many countries are engaged in international co-operation for AI, which is taking place in fora including the Trade and Technology Council (TTC), the Council of Europe (CoE), the EU, the G7 and G20, the Global Partnership on AI (GPAI), the Global Privacy Assembly (GPA), the Ibero-American Data Protection Network (RIPD), the Inter-American Development Bank (IDB), the International Telecommunications Union (ITU), the UN, UNESCO and the World Bank. Co-operation on AI research is also a priority.

Creating an evidence base to understand AI and emerging technologies

Timelier and better evidence is needed on AI-related trends and developments to understand the impact of technologies and design policies across domains, industries and Sustainable Development Goals (SDGs). Evidence should be reliable and representative and allow benchmarking. For example, gauging the diffusion and use of AI in a particular industry can help understand its impact on the quantity and quality of jobs and skill gaps in its workforce. Policymakers need timely data to stay apace of rapid developments in AI and ground policy in evidence. Interactive visualisations on OECD.AI show trends and developments by country and over time from different vantage points, from research, to jobs, skills, investment, open-source software development and education. The data confirm that AI development is booming, with 900% growth in open-source software development in only five years.[4] Social and gender gaps should also be considered. For example, the share of women authoring or co-authoring AI publications is below 20% in most countries.[5]

Better evidence-gathering tools[6] are needed to make informed decisions about how to govern AI, including what types or use cases to regulate, but also why, when and how to incentivise certain behaviours. Policymakers would benefit from evidence on the types, applications and characteristics of AI systems that have caused – or nearly caused – harm to people, societies, or the environment, such as risks that materialised into "AI incidents". Relevant characteristics include specifics about the data, model, task and context of AI systems involved. The OECD and partner organisations are developing a global *AI Incidents Tracker* to collect information from around the world. Tracking AI incidents would facilitate identifying AI applications that have caused harm, understanding their impacts and the causes of failure, to prevent harms from recurring and inform AI risk assessments and regulatory choices.

Developing and sharing tools for trustworthiness

In addition to evidence, AI practitioners and policymakers need tools and educational approaches to develop technologies consistent with OECD AI Principles. While tools are available, it is often difficult to find them and even more difficult to know which are most effective in each context. This could be resolved with a one-stop, open, interactive shop to compare tools that detect and remove biases from an AI system, or standards to promote AI accountability in a sector.

Implementing trustworthy AI and emerging technologies requires strong, structured and continuous engagement between companies and policymakers, given the rapid pace and technical complexity of AI developments. Corporate governance standards and sector-specific codes of conduct can be useful guidance. Private sector-led initiatives to embed trustworthy AI principles across business sectors and verticals are also valuable.

In addition, because AI knows no borders and many actors operate internationally, companies could benefit from government-endorsed international due diligence guidance to identify and address possible negative impacts that their operations and products could have. This includes engaging stakeholders from government, labour, affected communities and civil society. This guidance and related research can be part of a larger suite of tools for policymakers to support accountability for AI. The OECD is working to leverage the OECD Multinational Enterprise (MNE) Guidelines and accompanying Due Diligence Guidance for responsible business conduct for AI systems.

Leveraging the policy cycle to enable governance of emerging technologies

The policy cycle provides a starting point for the governance of emerging technologies. Building on the experience of more mature digital technologies such as AI, the policy cycle for emerging technologies includes four components (Figure 3):

- **Policy design**, including developing a national roadmap or strategy for the use and governance of a specific technology; conducting public consultations to raise awareness and promote social dialogue; designing a governance approach to co-ordinate policy implementation and oversight (e.g. assigning oversight to an existing ministry or creating an independent body); and exploring regulatory and non-regulatory approaches, going from voluntary frameworks, to soft law, regulatory experimentation and hard law or outright bans.
- **Policy implementation**, including identifying challenges and good practices in areas commonly outlined in internationally agreed principles[7], such as promoting inclusive growth and sustainable development; investing in R&D; fostering a digital ecosystem through competition and innovation; shaping an enabling policy environment, including via regulatory experimentation; and building human capacity to prepare for labour market transformation.
- **Policy evaluation**, including issuing regular reports highlighting milestones, accomplishments and lessons learned; establishing national observatories to monitor policy implementation; and establishing key performance indicators to measure progress with respect to specific targets (e.g. budget, skills, jobs, scholarships, research publications and patents, etc.)
- **International and multi-stakeholder co-operation**, including participating in and sharing good practices and lessons through fora like the OECD, the UN, the G7 and G20 and other international and intergovernmental organisations and multistakeholder groups. Cross-border co-operation in research is especially important for the governance of emerging technology, as is the involvement of national and international standardisation bodies to foster interoperability.

Figure 3. The emerging tech policy cycle

Source: Adapted from (OECD, 2021[35])

Conclusion: Multilateral co-operation to evolve with a changing context

AI is permeating economies, societies, governments and the environment, bringing tremendous opportunities but also risks. If appropriate safeguards are not in place, AI-enabled technologies, immersive tech and other next-generation computing technologies could inherit and widen the risks.

To harness the power of AI, policymakers can leverage and build upon the consensus-based OECD AI Principles, their definition of AI, the OECD framework for classifying AI systems and the Catalogue of Tools for Trustworthy AI.

Emerging technologies pose new challenges, but lessons can be learned from policy development processes designed for other technologies, including new approaches to regulatory experimentation. Technological innovations – such as immersive technologies – will continue to arise and impact our economies and societies, in both positive and not so positive ways. Research, sharing good practices and developing effective tools for accountability are crucial.

Several frameworks developed by the OECD can help policymakers evaluate risks and pilot policy and governance approaches before implementing them at scale. The OECD's longstanding expertise on digital technologies and its multi-stakeholder approach are well-placed to support countries building on the lessons to date in technology governance and applying them to emerging technologies. AI and emerging technologies are moving fast and so should governments. Policymakers must act now to ensure people and planet benefit from the opportunities at hand.

Notes

[1] OECD.AI uses data from LinkedIn to estimate the percentage of workers with AI skills by sector and gender. A joint publication by UNESCO, OECD and IDB (2022[49]) explores the effects of AI on the working lives of women.

[2] Several governments have included national security-related initiatives in their national AI strategies. These initiatives are not captured in this report and are left for further research.

[3] The draft request to CEN CENELEC is mentioned among others in CEN CENELEC news articles at https://www.cencenelec.eu/news-and-events/news/2022/newsletter/issue-34-etuc-s-position-on-the-draft-standardization-request-in-support-of-safe-and-trustworthy-ai/.

[4] From 2015 to 2020. Using data from GitHub.

[5] Using data from Scopus.

[6] In this context, “tools” is an umbrella terms that covers almost anything that helps make AI more trustworthy, from computer software and programming code to employee workshops and training or guidelines and standards.

[7] Including the OECD AI Principles, the OECD recommendation on Broadband Connectivity and the OECD recommendation on Enhancing Access to and Sharing of Data.

References

Ahmed, N. and M. Wahed (2020), "The De-democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research", arXiv:2010.15581 [cs]. [25]

Bertuzzi, L. (2022), *AI standards set for joint drafting among European standardisation bodies*, https://www.euractiv.com/section/digital/news/ai-standards-set-for-joint-drafting-among-european-standardisation-bodies/. [47]

Buolamwini, J. and T. Gebru (2018), "Gender shades: Intersectional accuracy disparities in commercial gender classification", *Proceedings of Machine Learning Research: Conference on fairness, accountability and transparency*, Vol. 81, pp. 1-15, http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf. [39]

Cassidy, M. (2017), "Who is liable if a self-driving car crashes? Tesla mishap raises issues", *Arizona Republic*, https://eu.usatoday.com/story/money/cars/2017/04/03/tesla-mishap-raises-issues-self-driving-liability/99880620/. [38]

Determann, L. (2018), "No One Owns Data"*, UC Hastings Research Paper*, No. 265, https://doi.org/10.2139/ssrn.3123957. [36]

European Investment Bank (2018), "A study on the deployment costs of the EU strategy on Connectivity for a European Gigabit Society", European Investment Bank, Kirchberg. [22]

G20 (2021), *Declaration of G20 Digital Ministers: Leveraging Digitalisation for a Resilient, Strong, Sustainable and Inclusive Recovery*, https://assets.innovazione.gov.it/1628084642-declaration-of-g20-digital-ministers-2021final.pdf. [44]

G20/OECD (2021), *G20 Guidelines for Financing and Fostering High-Quality*, http://www.g20.utoronto.ca/2021/G20-Guidelines-for-Financing-and-Fostering-High-Quality-Broadband-Connectivity-for-a-Digital-World.pdf. [24]

Ganguli, D. et al. (2022), "Predictability and Surprise in Large Generative Models", *arXiv:2202.07785 [cs]*. [46]

GSA (2022), *5G Standalone January 2022 – Member Report with Annex*, https://gsacom.com/paper/5g-standalone-january-2022-member-report-with-annex/. [37]

GSMA (n.d.), *Intelligent Connectivity. The fusion of 5G, AI and IoT*, https://www.gsma.com/ic/ (accessed on August 2022). [45]

Hansen, M. (2008), "Versatile, Immersive, Creative and Dynamic Virtual 3-D Healthcare Learning Environments: A Review of the Literature", *Journal of Medical Internet Research*, Vol. 10/3, https://doi.org/doi:10.2196/jmir.1051. [43]

Healthcare Insights (2022), *Immersive tech for healthcare*, https://thehealthcareinsights.com/immersive-tech-for-healthcare/ (accessed on June 2022). [9]

International Transport Forum (2021), *Artificial Intelligence in proactive road infrastructure safety management: Summary and conclusions*, OECD Publishing, Paris, https://www.itf-oecd.org/artificial-intelligence-proactive-road-infrastructure-safety-management. [5]

Khan, S., A. Mann and D. Peterson (2021), "The Semiconductor Supply Chain: Assessing National Competitiveness", Center for Security and Emerging Technology, https://doi.org/10.51593/20190016. [26]

Melchor-Couto, S. (2019), "Virtual worlds and language learning", *Journal of Gaming & Virtual Worlds*, Vol. 11/1, pp. 29-43, https://doi.org/10.1386/jgvw.11.1.29_1. [10]

Mohn, T. (2022), *Can A.I. all but end car crashes? The potential Is there.*, New York Times, https://www.nytimes.com/2022/04/19/technology/ai-road-car-safety.html (accessed on April). [4]

National Highway Traffic Safety Administration (2015), *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey*, US Department of Transportation, https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115. [6]

NIST (2022), *AI Risk Management Framework: Second Draft*, US National Institute of Standards and Technology, https://www.nist.gov/system/files/documents/2022/08/18/AI_RMF_2nd_draft.pdf. [34]

OECD (2022), "Broadband networks of the future", *OECD Digital Economy Papers*, No. 327, OECD Publishing, Paris, https://doi.org/10.1787/755e2d0c-en. [11]

OECD (2022), *OECD Broadband Portal, Database*, https://www.oecd.org/sti/broadband/broadband-statistics/. [21]

OECD (2022), "OECD Framework for the Classification of AI systems", *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, https://doi.org/10.1787/cb6d9eca-en. [12]

OECD (2022), "The AI footprint: measuring the environmental impacts of AI compute and applications", *OECD Digital Economy Papers*, OECD Publishing, Paris, https://oecd.ai/en/footprint. [29]

OECD (2021), "Good Practice Principles for Data Ethics in the Public Sector - OECD", OECD Publishing, Paris, https://www.oecd.org/gov/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm. [32]

OECD (2021), *Issues report, Meeting of the OECD Council - Paris, 5-6 October*, OECD Publishing, Paris, https://www.oecd.org/mcm/MCM_2021_Part_2_[CMIN_2021_15_EN].pdf. [28]

OECD (2021), "Promoting high-quality broadband networks in G20 countries", OECD Publishing, Paris, https://doi.org/10.1787/cf0093dc-en. [23]

OECD (2021), *Recommendation of the Council on Broadband Connectivity*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0322. [15]

OECD (2021), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463. [14]

OECD (2021), "State of implementation of the OECD AI Principles: Insights from national AI policies", *OECD Digital Economy Papers*, No. 311, OECD Publishing, Paris, https://doi.org/10.1787/1cd40c44-en. [35]

OECD (2020), "Open, Useful and Re-usable data (OURdata) Index: 2019", *OECD Policy Papers on Public Governance*, No. 1, OECD Publishing, Paris, https://www.oecd.org/gov/digital-government/policy-paper-ourdata-index-2019.htm. [31]

OECD (2020), "Smart Cities and Inclusive Growth. Building on the outcomes of the 1st OECD Roundtable on Smart Cities and Inclusive Growth", OECD Publishing, Paris, https://www.oecd.org/cfe/cities/OECD_Policy_Paper_Smart_Cities_and_Inclusive_Growth.pdf. [1]

OECD (2019), *Artificial Intelligence in Society*, OECD publishing, Paris, https://doi.org/10.1787/eedfee77-en. [13]

OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, https://doi.org/10.1787/276aaca8-en. [48]

OECD (2019), *Hello, World: Artificial Intelligence and its use in the public sector*, OECD Publishing, Paris, https://oecd-opsi.org/publications/hello-world-ai/. [2]

OECD (2019), "Measuring distortions in international markets: The semiconductor value chain", *OECD Trade Policy Papers*, No. 234, OECD Publishing, Paris, https://doi.org/10.1787/8fe4491d-en. [27]

OECD (2019), *Recommendation of the Council on Artificial Intelligence*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. [16]

OECD (2019), "Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)", *OECD Digital Economy Papers*, No. 291, OECD Publishing, Paris, https://doi.org/10.1787/d62f618a-en. [30]

OECD (2013), *Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188. [17]

OECD (forthcoming), "Going Digital Guide to Data Governance Policy Making", OECD Publishing, Paris. [19]

OECD (forthcoming), "Going Digital to Advance Data Governance for Growth and Well-being", OECD Publishing, Paris. [18]

OECD (forthcoming), "Measuring national compute capacity for Artificial Intelligence (AI). Existing measurement tools and preliminary findings", OECD Publishing, Paris. [20]

Raji, I. and J. Buolamwini (2019), "Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products", *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 429-435, https://www.thetalkingmachines.com/sites/default/files/2019-02/aies-19_paper_223.pdf. [40]

Salvi, A., P. Wyckoff and A. Vourc'h (2022), "Using Artificial Intelligence in the workplace: What are the main ethical risks?", *OECD Social, Employment and Migration Working Papers*, Vol. 273, https://doi.org/10.1787/840a2d9f-en. [33]

Slater, M. (2009), "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments", *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 364, pp. 3549-3557. [42]

Soliman, M., J. Peetz and M. Davydenko (2017), "The impact of immersive technology on nature relatedness and pro-environmental behavior", *Journal of Media Psychology*, Vol. 29, pp. 8-17. [41]

Stephenson, N. (1992), *Snow Crash*, Spectra. [7]

UNESCO/OECD/IDB (2022), *The Effects of AI on the Working Lives of Women*, https://doi.org/10.1787/14e9b92c-en. [49]

Verizon (2022), *5G-powered digital twin*, https://www.verizon.com/business/resources/5g/5g-business-use-cases/business-intelligence/digital-twin/ (accessed on June 2022). [8]

World Health Organisation (2018), *Global status report on road safety*, https://www.who.int/publications/i/item/9789241565684. [3]