



Economic and Social Council

Distr.: General
29 January 2024

Original: English

Committee of Experts on Public Administration

Twenty-third session

New York, 15–19 April 2024

Item 9 of the provisional agenda*

Digital government

Artificial intelligence governance to reinforce the 2030 Agenda and leave no one behind

Note by the Secretariat

The Secretariat has the honour to transmit to the Committee of Experts on Public Administration the paper prepared by Committee members Sherifa Sherif and Carlos Santiso.

* E/C.16/2024/1.



Artificial intelligence governance to reinforce the 2030 Agenda and leave no one behind

Summary

Artificial intelligence is increasingly being used worldwide with immense potential benefits, augmenting human capabilities, increasing the well-being of people and contributing to the betterment of society. However, as the technology continues to evolve at an unprecedented pace, many challenges, risks and ethical concerns remain that must be addressed urgently. Governments have a particularly important role to play, both as regulator and user of artificial intelligence, especially given their huge influence over people's lives.

The authors elaborate on the evolution of artificial intelligence, noting that it has transformed industries and the daily lives of people in ways that were once unimaginable. They also provide an overview of the potential of artificial intelligence for accelerated implementation of the 2030 Agenda for Sustainable Development, while giving a snapshot of its benefits for selected Sustainable Development Goals.

The authors then elaborate on the multitude of risks and challenges associated with using artificial intelligence in the long run, especially in developing countries, raising concerns about ethics, equity, transparency and compliance with existing and emerging regulations.

The authors call for artificial intelligence governance aimed at ensuring equal access to the benefits of artificial intelligence, protecting digital rights and preventing harm. They provide an overview of existing artificial intelligence governance practices while highlighting the main elements required to build a strong artificial intelligence governance framework.

In closing, the authors emphasize the need to continue the ongoing global discussion and build a shared understanding of both the positive and negative impacts of artificial intelligence. Principles will have to be identified, policy priorities set and policy coherence ensured so that artificial intelligence can reinforce the 2030 Agenda and leave no one behind.

I. Introduction

1. In today's technologically advanced world, artificial intelligence is increasingly used by governments, organizations and people with immense potential benefits. As artificial intelligence systems become more capable and integrated into our daily lives, they can increasingly augment human capabilities, increase the well-being of people and contribute to the betterment of society. Artificial intelligence can contribute to sustainable economic growth, increase innovation and productivity and help to respond to global challenges. For instance, artificial intelligence-driven climate modelling could help to address the pressing issue of climate change, while artificial intelligence in education could enable personalized learning experiences and make education more accessible to those in remote or underserved areas.

2. Governments have a particularly important role as both regulator and user of artificial intelligence, especially given their impact on people's lives and responsibility to ensure their well-being.¹ They have started to invest in artificial intelligence technologies to support smart policymaking and operational efficiency with a view to solving various public sector problems, such as traffic flows, access to education and health-care services, infrastructure monitoring, cyberattacks and much more. Currently, artificial intelligence is used in different areas of the public sector, such as law enforcement, judicial administration, fraud prevention, taxation and the fight against corruption.² Artificial intelligence can be a fantastic booster of innovation in government, altering how public administrations think and operate, with the potential to improve internal operational efficiency, policymaking effectiveness (including better targeting of public spending and social transfers), public service delivery and responsiveness (including a move towards more agile, personalized, proactive and people-centred public services), government integrity and fraud prevention, and the evaluation of public policies.

3. However, many challenges and ethical concerns remain that must be addressed urgently as the technology continues to evolve. Privacy and security issues and the potential misuse of artificial intelligence require careful consideration and regulation. A particular challenge concerns the ways in which artificial intelligence is changing public governance and its responsible use in and by the public sector itself, which has a special duty when it comes to protecting people's digital rights. Robust guard rails are required to govern the deployment of artificial intelligence in the public sector, in particular in sensitive policy domains, such as welfare benefits, fraud detection, law enforcement and judicial sentencing, and personalized services. Governing with artificial intelligence is set to be a defining challenge of the digital era.

4. Going forward, there is a need for a continued global dialogue and the building of a shared understanding of both the positive and negative impacts of artificial intelligence on the machinery of government. Principles will have to be identified, policy priorities set and policy coherence ensured to reinforce achievement of the 2030 Agenda for Sustainable Development. Collective action towards artificial intelligence governance is important to minimizing its negative consequences, with a

¹ See for example Organisation for Economic Co-operation and Development (OECD) and Development Bank of Latin America and the Caribbean, *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector in Latin America and the Caribbean* (Paris, OECD, 2022); and Jamie Berryhill and others, "Hello, world: artificial intelligence and its use in the public sector", OECD Working Papers on Public Governance, No. 36 (OECD, 2019).

² See for example World Economic Forum, "Hacking corruption in the digital era: how tech is shaping the future of integrity in times of crisis", May 2020; and Carlos Santiso, "Trust with integrity: harnessing the integrity dividends of digital government for reducing corruption in developing countries", DESA Working Paper, No. 176 (New York, United Nations, Department of Economic and Social Affairs, 2022).

focus on leaving no one – and no country – behind, and to ensuring equal access to technology. The United Nations has a critical and pivotal role to play in fostering a human-centred, rights-based deployment of artificial intelligence and supporting the establishment of much-needed global governance and regulation of artificial intelligence.

5. The present paper builds on the previous work of the Committee on digital government. At its twenty-second session, the Committee noted that, going forward, efforts were urgently needed to, among other things, address digital divides, improve data governance and mitigate risks arising from the use of new technologies such as artificial intelligence and social media, by adopting new policy and regulatory regimes and standards.

6. Member States, in the political declaration of the high-level political forum on sustainable development convened under the auspices of the General Assembly, in September 2023 (resolution 78/1), agreed that they would seek to better realize the benefits and address the challenges of artificial intelligence.

7. In 2023, the Secretary-General announced the creation of a new advisory body to assess the risks, opportunities and international governance of artificial intelligence in support of the efforts of the international community to govern artificial intelligence. The High-level Advisory Body on Artificial Intelligence, in its first interim report, confirmed that global artificial intelligence governance was essential to reap the significant opportunities of artificial intelligence, while mitigating its risks for States, communities and individuals today and in the future. In his report entitled “Our Common Agenda” (A/75/982), the Secretary-General also suggested the establishment of a “global digital compact” to be agreed at the Summit of the Future, in 2024, which could, among other things, promote the regulation of artificial intelligence to ensure that it is aligned with shared global values.³

II. Definition and evolution of artificial intelligence

8. Currently there is no universally agreed definition of artificial intelligence. For the purposes of the present paper, artificial intelligence is defined as the ability of machines to imitate intelligent human behaviour by performing various cognitive tasks, such as sensing, processing oral language, reasoning, learning, making decisions and demonstrating an ability to manipulate objects accordingly.⁴ Artificial intelligence is essentially using algorithms to imitate the operations and procedures of the human brain with the aim of making computers think and act like humans. The primary functions of these algorithms are pattern recognition, prediction and control, making artificial intelligence a key issue on the agenda of governments and regional and international organizations. Artificial intelligence has the ability to process and analyse vast amounts of data and can be harnessed to support everyday tasks, as well as to tackle complex problems that have long eluded human understanding.

Traditional versus generative artificial intelligence

9. Traditional artificial intelligence receives input and produces an output, with data being analysed and used in decision-making and forecasting. It is still very popular and used to power a large population of artificial intelligence systems, such as chatbots and predictive analytics. Traditional artificial intelligence relies on rule-based approaches, whereby explicit instructions and predefined rules are programmed

³ For more information, see <https://www.un.org/techenvoy/global-digital-compact>.

⁴ See also Economic and Social Commission for Asia and the Pacific, “Artificial intelligence in Asia and the Pacific”, policy brief, November 2017.

to enable the system to perform specific tasks and generate outputs. These rules are designed by human experts on the basis of their understanding of the problem domain.

10. On the other hand, generative artificial intelligence takes a data-driven approach, learning patterns and structures from large data sets using machine learning techniques. Instead of relying on explicit rules, generative artificial intelligence models learn from the data and generate new content by capturing underlying patterns and relationships within the data. Generative artificial intelligence offers opportunities for users to be more creative and innovative, thus reducing the amount of time spent on the ideation process. ChatGPT is an example of a generative artificial intelligence tool.

Evolution

11. The evolution of artificial intelligence has been a remarkable journey, with countless breakthroughs and innovations accelerating and propelling the field forward. From its humble beginnings in the 1950s to the sophisticated deep learning models seen today, artificial intelligence has transformed industries and the daily lives of people in ways that were once unimaginable.

12. Artificial intelligence was introduced in the 1950s with the work of Alan Turing, who invented the Turing test to determine whether a machine could mimic human intelligence. In the 1960s, the first artificial intelligence programming language, LISP, was developed by John McCarthy. Early artificial intelligence systems focused on symbolic reasoning and rule-based systems, which eventually led to the development of expert systems in the 1970s and 1980s – computer systems emulating the decision-making ability of human experts.

13. In the 1990s, the focus of artificial intelligence was shifted towards machine learning and data-driven approaches, a result of the increased availability of digital data and the advancement of computers. A rise in neural networks allowed artificial intelligence systems to learn from data, which led to better performance and adaptability. In the 2000s, artificial intelligence research delved into new areas, including natural language processing, computer vision and robotics, which paved the way for today's artificial intelligence revolution.

14. Today, global government expenditure in artificial intelligence is increasing, particularly in Canada, China, the United Kingdom of Great Britain and Northern Ireland and the United States of America. In 2020, the Government of the United States provided more than \$1 billion in funding for artificial intelligence projects. In March 2021, the Government of Canada pledged to spend more than \$500 million to promote its artificial intelligence initiatives. Following the release of ChatGPT, the market also saw an incredible influx of hundreds of artificial intelligence products.

15. The main impacts of artificial intelligence are expected in the following sectors: health care; automotive; financial services; retail and consumer; technology, communications and entertainment; manufacturing; energy; and transport and logistics.⁵ In the automotive industry, for example, artificial intelligence can be used to improve vehicle performance, driver safety and passenger experience, while artificial intelligence robots are being used in assembly lines. In the health-care sector, artificial intelligence can help to decrease human errors, assist medical professionals, and provide patient services on a 24/7 basis.

16. The speed with which artificial intelligence is evolving is unprecedented. This is partly because the technology can augment itself, improving its own capabilities. Artificial intelligence also feeds on ever-expanding sources of new data in the context

⁵ Ibid.

of the so-called “data revolution”.⁶ As more and more data are generated from an increased use of new technologies such as the Internet of things and unstructured big data, and as the Internet gets faster, the rate of development of artificial intelligence only increases. In addition, artificial intelligence is a general-purpose technology with a limitless scope of applications. While most people are more familiar with ChatGPT or text-to-image generators, artificial intelligence can also be used in critical infrastructure, such as energy systems or water supplies, making its responsible management and guard rails indispensable. An important risk highlighted at the recent AI Safety Summit 2023 in the United Kingdom is that generative artificial intelligence might run amok, operating beyond human anticipation or intended oversight and making decisions or taking actions that the developers of the technology did not foresee or intend, with potentially devastating consequences.

17. As the boundaries of what artificial intelligence can achieve are continuously being pushed, new challenges and ethical dilemmas will inevitably appear. However, by fostering a collaborative environment between different actors, such as governments, international organizations, businesses and researchers, the evolution of artificial intelligence can be guided by a collective vision that prioritizes the betterment of society and the well-being of individuals worldwide. With the deadline for achieving the 2030 Agenda fast approaching, and moving forward, continuous investments in research and development should take place, ensuring the responsible and ethical harnessing of the potential of artificial intelligence to address global challenges and create a better world for all, leaving no one – and no country – behind.

III. Potential of artificial intelligence for accelerating implementation of the Sustainable Development Goals

18. Artificial intelligence can enhance creativity and problem-solving and holds the promise of opening new avenues in designing sustainable solutions to accelerate implementation of the 2030 Agenda, fostering innovation across different industries and sectors. The present section gives a snapshot of the potential benefits of artificial intelligence in support of accelerated implementation of selected Sustainable Development Goals, while keeping in mind that all of them are interlinked.⁷

Goal 1 (poverty eradication)

19. The 2030 Agenda is aimed at ensuring that no one is being left behind by ending poverty in all its forms everywhere. Artificial intelligence can support this pledge, among others, by:

- **Supporting the identification of vulnerable groups and tracking of poverty levels.** This allows for better targeting of pro-poor and equity-based policies and programmes and ensures that aid and resources are being allocated where they are needed the most. In addition, upcoming trends and needs can be predicted, allowing for better planning of future interventions.
- **Improving access to basic services.** Artificial intelligence can improve access of the poor and vulnerable groups to basic services, such as education and health care. For example, it can review vast amounts of health-care data and demographic information to identify regions in need of hospitals, mobile health units or telemedicine services.

⁶ Independent Expert Advisory Group on the Data Revolution for Sustainable Development, “A world that counts: mobilizing the data revolution for sustainable development”, November 2014.

⁷ See also Ricardo Vinuesa and others, “The role of artificial intelligence in achieving the Sustainable Development Goals”, *Nature Communications*, vol. 11, 2020.

- **Fostering agricultural development.** Artificial intelligence can, for example, predict crop yields and prices, allowing farmers to obtain maximum profit, while also monitoring the health of crops and soil. In addition, artificial intelligence robots can harvest higher volumes of crops at a faster pace, increasing the income of farmers.
- **Promoting financial inclusion.** Artificial intelligence can facilitate financial inclusion by offering affordable and accessible banking services to disadvantaged population groups. It can assess creditworthiness more accurately, making it easier for individuals and small businesses to obtain access to loans and financial services. Artificial intelligence can assist in the efficient distribution of microfinance by identifying the most deserving recipients on the basis of data analysis.

Goal 4 (quality education)

20. Artificial intelligence has the potential to address some of the most pressing challenges in the education sector today, transforming teaching and learning, creating a more personalized, effective and accessible education system for all and speeding up the achievement of Goal 4. Positive potential impacts of artificial intelligence on the education sector include:

- **Personalized and dynamic learning.** Artificial intelligence can tailor educational content to individual students, thus enhancing learning outcomes and student engagement. It can also lead to the creation of interactive educational tools.
- **Inclusion.** With advanced translation and comprehension abilities, artificial intelligence technology can help to bridge knowledge and language gaps, making quality education and resources accessible to a much broader audience irrespective of language or local/regional constraints. Generative artificial intelligence-driven platforms can provide round-the-clock assistance to learners, making quality education more universally available. In addition, artificial intelligence-powered tools can support students with disabilities (e.g. in speech-to-text conversion). This is important, as leaving no one behind must mean that everyone will have access to and benefit from the current technological revolution.

Goal 8 (decent work and economic growth)

21. Artificial intelligence is becoming an indispensable tool for economic activities, and its tremendous contribution to economic development is rapidly becoming evident, with obvious intersections between artificial intelligence and various sectors of economic activity. For instance, artificial intelligence can improve production, increase efficiency and improve safety along production lines while lowering cost, thus allowing different sectors to provide good services at competitive prices.

22. Artificial intelligence algorithms are data-driven and able to learn data trends over time, making them highly suitable in forecasting economic indicators, such as growth, interest, exchange and inflation rates, which are vital for monetary policy management and economic stability. Accurate predictions of these indicators can support policymakers and make them more proactive in predicting upcoming challenges, such as the next financial crisis. With regard to trading assets, such as stocks and bonds, artificial intelligence technology can also predict price movements, allowing policymakers to trade during optimal times.

23. At the global level, generative artificial intelligence can bridge language and knowledge gaps, fostering greater international collaboration. At the national level,

artificial intelligence can foster decent work and economic growth, among others, through:

- **Increased productivity and assistance with complex tasks.** Generative artificial intelligence can assist humans in performing complex tasks, improving output and efficiency. It can manage, analyse and process information more efficiently than humans, providing the potential to increase overall productivity. Estimates suggest that the impact of generative artificial intelligence on productivity could add up to \$15.7 trillion in value to the global economy by 2030.⁸ The ability of artificial intelligence to save millions of work-hours is also a key justification for its use in government procedures. As a result, employees are freed up to concentrate on duties that are more intellectually crucial. Recent estimates conclude that automating tasks of government workers can save from \$3.3 billion to \$41.1 billion annually.⁹
- **Cost reduction and efficiency gains.** Artificial intelligence and automation can speed up processing, cut costs and lead to faster provision of services. Artificial intelligence-powered robots can also work round the clock, ensuring a continuous supply of output. Some 33 per cent of manufacturers are already reducing their labour costs through artificial intelligence and other technologies. Artificial intelligence and automation are also expected to reduce unplanned downtime and product defects by up to 50 per cent and to increase manufacturing efficiency by up to 20 per cent.¹⁰
- **Creation of new jobs.** With the evolution of generative artificial intelligence, new sectors and professions may emerge, similar to how the information technology revolution led to the creation of a plethora of technology jobs. The World Economic Forum, in its *The Future of Jobs Report 2020*, estimates that, by 2025, around 97 million new employments will have been generated by the development of artificial intelligence and technology.
- **Improved access to expertise and services.** In areas with a shortage of experts, generative artificial intelligence can provide essential expertise and services.

24. Artificial intelligence-related activities will be the driving force for further economic development in many countries and will result in fundamental shifts in production structures and approaches and in the quantity and quality of consumption. With the rapidly evolving advancements in the technological landscape, artificial intelligence stands to reshape economies, labour markets and industries, revolutionizing different sectors. Its implications on the global workforce and economic disparities will require well-thought-out policies.

Goal 9 (industry, innovation and infrastructure)

25. Artificial intelligence has the potential to accelerate the achievement of Goal 9, among others, through:

- **Innovation acceleration.** Generative artificial intelligence can accelerate research and development in various industries by analysing vast amounts of data, predicting outcomes and generating innovative solutions.

⁸ PricewaterhouseCoopers, “Sizing the prize: what’s the real value of AI for your business and how can you capitalize?”, 2017; see also Michael Chui and others, *The Economic Potential of Generative AI: The Next Productivity Frontier* (McKinsey & Company, 2023).

⁹ Deloitte, “AI-augmented government: using cognitive technologies to redesign public sector work”, 2017.

¹⁰ Saxon, “Impact of AI in manufacturing: improved quality and efficiency”, December 2022.

- **Infrastructure management.** Generative artificial intelligence can enhance infrastructure management by predicting potential infrastructure failures, optimizing traffic systems in real time and managing large-scale utilities more efficiently.
- **Enhanced manufacturing.** Advanced artificial intelligence systems can drive automation, optimize supply chains and predict machinery maintenance issues, increasing overall efficiency in the manufacturing sector.

26. Government should strive to introduce and promote new technologies, facilitate international trade and enable the efficient use of resources, as well as scale up investment in scientific research and innovation.

Goal 13 (climate action)

27. Artificial intelligence can help to combat climate change by supporting mitigation and adaptation measures, among others, through the following:

- **Improving modelling and predictions of climate change patterns.** Artificial intelligence can help communities and authorities to draft more effective adaptation and mitigation strategies and make them better prepared for upcoming extreme weather events, such as heatwaves, droughts and floods, in line with the Secretary-General's Early Warnings for All initiative.
- **Improving urban planning and traffic management.** Artificial intelligence can decrease greenhouse gas emissions and make cities more sustainable and liveable. It can also track pollution levels, enabling local governments to alert the public of dangerous levels.
- **Supporting carbon neutrality.** Artificial intelligence has a key role to play in supporting countries on their path to carbon neutrality. For example, it can optimize manufacturing processes, thus reducing their environmental impacts, reduce traffic and increase the efficiency of renewable energy sources.

IV. Risks and challenges of artificial intelligence

28. Although the widespread application of artificial intelligence is likely to cause a short-lived economic boost and holds a multitude of potential benefits, in the long run, overreliance on artificial intelligence is likely to pose a multitude of risks and challenges.

Overdependence on technology

29. Overreliance on artificial intelligence could reduce human interaction and connection, critical thinking and other essential soft skills, leading to loss of creativity, social skills and empathy. In addition, technical difficulties and glitches can disrupt education, learning and productivity.

Job displacement and changing requirements

30. The most significant concern regarding artificial intelligence is involuntary job loss and, more generally, the future of work in the age of artificial intelligence. Artificial intelligence has the potential to eliminate jobs by imitating human cognitive processes and carrying out, at much quicker rates and with reduced operational expenses, several routine activities presently done by employees.

31. Manual labour and jobs that involve repetitive tasks or can be systematized are most at risk of being automated, including not only blue-collar jobs but also certain

white-collar professions (e.g. accountants, editors, retail and courier services workers, security personnel and even physicians). The proliferation of artificial intelligence could also contribute to the erosion of middle-wage job opportunities, particularly in sectors reliant on human creativity and content generation. According to the World Economic Forum *The Future of Jobs Report 2020*, around 85 million jobs may be lost by 2025 as a result of the development of artificial intelligence and related technologies.

32. As the value of certain forms of human labour decreases as a result of the capabilities of generative artificial intelligence, there might be a turbulent phase of adjustment. Traditional roles might experience diminished demand and compensation, widening disparities within the workforce of many countries and potentially exacerbating international inequalities. In addition, the demand for new skills might grow, paired with a shift towards more flexible work arrangements. Governments might need to intervene more proactively, for example by establishing new labour policies, supporting the transitioning of workers to new industries or even exploring new concepts, such as universal basic income, in response to drastically reduced labour costs. They will also have to build the internal capacities of the public sector workforce so that not all artificial intelligence-based technology development is being outsourced to private sector partners and they can better understand, develop and govern these technologies.

Lack of skills

33. Artificial intelligence is also predicted to have a favourable overall influence on the employment market, generating plenty of opportunities for competent workers despite the loss of some occupations. That positive influence of artificial intelligence is feasible, however, only if countries reskill and upskill their workforce with the needed skills and competencies. IBM estimates that around 40 per cent of workers (1.4 billion of the 3.4 billion in the global workforce) will need to reskill over the next three years.¹¹ Digital illiteracy will also need to be addressed.

Loss of traditional industries

34. The use of artificial intelligence could lead to the loss of traditional industries. In developing countries with economies reliant on traditional industries, rapid automation driven by generative artificial intelligence could lead to economic instability.

Lack of quality data

35. Artificial intelligence programmes are only as good as the underlying algorithms and data on which they operate. The lack of finely tuned and ethically reviewed algorithms, using reliable registries with accurate and representative data, can have a negative impact on results and increase the risk of bias, possibly leading to new forms of exclusion and discrimination. This is particularly an issue in developing, “data-poor” countries, which lack high-quality data in relation to their citizens. Accordingly, the Organisation for Economic Co-operation and Development (OECD) recommends paying greater attention to the ethics of data, in particular in the public sector.¹² Artificial intelligence strategies should be further integrated into or linked with data governance and infrastructure strategies and efficient, digitalized administrative registries. However, this remains a pervasive challenge.

¹¹ IBM Institute for Business Value, “Augmented work for an automated, AI-driven world: boost performance with human-machine partnerships”, 2023.

¹² OECD Good Practice Principles for Data Ethics in the Public Sector.

Economic disparities and equity issues

36. Advancements in digital technology are astounding, with increased use of digital tools and improvements in the Internet infrastructure, leading to noticeable societal effects. However, usage and development of artificial intelligence vary widely among countries, industries and sectors, and social systems. The advantages of the digital revolution need to be evenly distributed across economies and individuals by increasing access to digital possibilities and closing the digital divide.

37. If generative artificial intelligence technologies are developed and owned primarily by a handful of countries or corporations, it could lead to significant economic imbalances globally and an even wider economic divide between countries. The lack of access to new technologies in developing countries threatens to increase inequalities between countries. Manufacturing giants equipped with artificial intelligence technologies will experience accelerated growth, leaving behind developing countries without access to such advancements. Along the same lines, if only certain regions or groups have access to advanced artificial intelligence educational tools, it could widen educational disparities within and between countries. If left unchecked, these disruptions caused by advanced technologies may have enormous societal consequences. It is therefore important to ensure that all countries and people benefit from artificial intelligence development.

38. Governments should strive to minimize economic polarization to prevent inequitable distribution of the benefits of artificial intelligence, as well as a widening economic divide with a larger proportion of wealth being controlled by those who own and manage the generative artificial intelligence systems.

Moral and ethical issues

39. While new technologies can be economically beneficial, they can also trigger adversity for certain communities by fostering bias and discrimination. This dichotomy necessitates astute policy interventions. Machine learning biases, for example, particularly in terms of racial profiling, can incorrectly identify basic information about users, which can result in unfair denial of access to health care and loans or mislead law enforcement in identifying criminal suspects. However, instilling moral and ethical values in artificial intelligence systems, especially for decision-making processes with significant consequences, remains a considerable challenge.

40. Nevertheless, a number of countries are working on or have already established artificial intelligence governance procedures and rules to minimize the emergence of prejudice or discrimination in algorithms, for example by increasing their transparency and establishing open public algorithm registers. Colombia, for example, was one of the first OECD countries to adopt an ethical framework for its artificial intelligence strategy.

Misinformation and manipulation

41. Content generated by artificial intelligence, such as “deepfakes”,¹³ increasingly contributes to the spread of false information and the manipulation of public opinion. Efforts to detect and combat such misinformation are critical, as it may decrease the perceived legitimacy of public institutions, deepen political polarization and favour populist movements.

¹³ Video of a person in which the face or body has been digitally altered so that the person appears to be someone else, typically used maliciously or to spread false information.

Privacy issues and security risks

42. As generative artificial intelligence systems require vast amounts of data, there are concerns as to who controls the data and regarding potential data monopolies or misuse. The potential misuse of artificial intelligence (e.g. cyberattacks) needs to be prevented and concerns about artificial intelligence surveillance, namely the use of artificial intelligence technology to monitor and analyse human behaviour for security, law enforcement and marketing purposes, addressed. Strict data protection regulations and safe data handling practices are required. Global norms and regulations protecting against artificial intelligence security threats are also required, especially in view of rising concerns over the use of artificial intelligence-driven autonomous weaponry by rogue States or non-State actors.

Lack of transparency, unintended consequences, and potentially existential risks

43. The lack of transparency in artificial intelligence systems, particularly in deep learning models, which are complex and difficult to interpret, needs to be addressed urgently, as there might be distrust and resistance to adopting this technology if people cannot comprehend how an artificial intelligence system comes to conclusions or solutions. In addition, when an algorithm is a “black box”, it is very difficult to oversee it effectively.

44. As a result of their complexity and lack of human oversight, artificial intelligence systems might also exhibit unexpected behaviour or make decisions with unforeseen consequences, with a negative impact on individuals or society as a whole. In addition, the possible development of artificial general intelligence (self-teaching and able to carry out a wide range of tasks autonomously) that surpasses human intelligence is starting to raise concerns due to possible unintended and potentially catastrophic consequences if these advanced artificial intelligence systems do not align with human values or priorities.¹⁴

V. Artificial intelligence governance to ensure sustainable development and leave no one behind

45. The rapid worldwide adoption of artificial intelligence raises different concerns about ethics, equity, transparency and compliance with other regulations. Without proper governance, artificial intelligence systems could pose immense risks as illustrated in section IV, especially for developing countries.

46. Governments need to appraise and take stock of artificial intelligence technologies, acknowledging both their potential benefits and inherent risks. Short-term challenges, if not addressed, could grow into long-term systemic issues. Therefore, there is an immediate need for a comprehensive rethinking and redesigning of policies, social security systems, labour markets and taxation frameworks, while ensuring transparency, accountability and human oversight, and respect for shared norms and values, such as those enshrined in the Charter of the United Nations, the Universal Declaration of Human Rights and international law.¹⁵

47. The responsible and ethical use of artificial intelligence in the public sector deserves further consideration. Many digitally advanced OECD countries that were early adopters of the use of artificial intelligence in government processes and services are paying greater attention to the benefits and specific challenges and risks

¹⁴ See for example Bernard Marr, “The 15 biggest risks of artificial intelligence”, Forbes, 2023.

¹⁵ See also the 2023 interim report of the High-level Advisory Body on Artificial Intelligence.

of artificial intelligence for the public sector, not only as a regulator but also as a critical user of artificial intelligence.

Existing artificial intelligence governance practices

48. The creation and execution of national artificial intelligence policies and strategies are at various stages in different countries.¹⁶ Some countries, including Canada and Finland, started to develop their national artificial intelligence plans as early as 2017, followed by France, Germany, Japan and the United Kingdom in 2018. Other countries have more recently introduced national artificial intelligence strategies, including Brazil, Egypt, Hungary, Poland and Spain. Today, more than 60 countries have a dedicated artificial intelligence strategy,¹⁷ and artificial intelligence policy development and consultation procedures are under way in several other countries.

49. In recent years, an increasing number of Governments have also introduced specific policies and standards aimed at governing artificial intelligence in the public sector. While highly decentralized in its deployment, efficient governance of artificial intelligence in the public sector requires a strong steer and oversight from the centre of government to ensure uniform rules and consistent standards. The Directive on Automated Decision Systems of Canada, for example, describes how the Government uses artificial intelligence to guide decisions in several departments. A scoring system is being used to assess the kind of human intervention, peer review, monitoring and contingency planning needed to achieve an artificial intelligence tool built to serve its citizens.

50. Many countries, such as Chile and France, are mandating transparency in the artificial intelligence algorithms used by public sector entities through open registers of public algorithms and by enforcing access-to-information rules on algorithms used by government entities. Chile is also working on establishing the first regulation on algorithmic transparency in the public sector in Latin America and the Caribbean. Several cities, including Amsterdam, Barcelona and Helsinki, have also established open registries. In addition, the European Commission created the European Centre for Algorithmic Transparency.

51. Several other countries, including the United Kingdom and the United States, are using government procurement rules to embed core (ethical) principles in the procurement of artificial intelligence solutions for public sector entities. By insisting on certain standards for contractors, they can set an example that will influence the behaviour of the wider marketplace.

52. Several countries and regions also aim to protect digital rights through regulations and policies. Spain, for example, adopted the Charter of Digital Rights, and the European Union signed the Declaration on Digital Rights and Principles.

53. With regard to the use of artificial intelligence in the private sector, several countries in Asia and the Pacific released a number of policies and regulations to tackle ethical concerns (e.g. the artificial intelligence governance and ethics initiatives of Singapore). Others, including the United Kingdom and the United States, building on the approach of the Group of Seven to artificial intelligence regulation, are promoting the use by the private sector of voluntary codes of conduct.

54. Some countries are also envisaging the establishment of artificial intelligence agencies, similar to those established on data protection, which could lead detailed

¹⁶ See also the OECD live repository of national policies and strategies on artificial intelligence, available at <https://oecd.ai/en/dashboards/overview>.

¹⁷ Carlos Santiso, “Public governance in the age of artificial intelligence”, *Governance Matters* (Chandler Institute of Governance, 2023).

impact assessments, test potential solutions and undertake research on their potential positive and negative impacts before they are rolled out. Spain, for example, created Europe's first artificial intelligence supervisory agency in 2023.

55. In addition, there are regional initiatives aimed at regulatory convergence regarding artificial intelligence, including in Africa, Asia and Europe. In December 2023, for example, the European Union agreed on its Artificial Intelligence Act, which will come into effect in 2026. Pursuant to the Act, different artificial intelligence systems are classified according to the risk that they pose to users, with different levels of regulation. The Act, among other things, prohibits artificial intelligence systems that pose an "unacceptable risk" from being deployed in the European Union and imposes different levels of obligations on artificial intelligence systems categorized as "high risk" or "limited risk" (e.g. deepfakes). In October 2023, 20 countries adopted the Santiago Declaration to Promote Ethical Artificial Intelligence in Latin America and the Caribbean, with support from the United Nations Educational, Scientific and Cultural Organization and the Development Bank of Latin America and the Caribbean.

Going forward

56. The interest in artificial intelligence governance is growing, with a focus on how much of our daily life should be shaped by algorithms and identifying who is in control of monitoring them. A global and united approach to artificial intelligence governance should prevent regulatory fragmentation and allow for constructive use of artificial intelligence technologies while ensuring equal access, protecting human rights ("digital rights") and preventing harm.¹⁸ This is particularly important, as the technology space operates across borders, making international coordination and cooperation indispensable. It is also in line with the proposed global digital compact to be adopted at the Summit of the Future in 2024, which is intended to outline shared principles for an open, free and secure digital future for all. Initiatives such as the Global Partnership on Artificial Intelligence and the OECD Artificial Intelligence Policy Observatory can support the required information exchange, dialogue and collaboration between countries and different stakeholders.

57. The following are elements for which artificial intelligence governance should aim:

- Creating institutional and legal frameworks for the application of artificial intelligence technology.
- Complying with data governance rules and privacy regulations; outlining guidelines for access to and management of personal data.
- Addressing moral, ethical and security issues linked with artificial intelligence.
- Preventing misinformation and manipulation.
- Fostering safety, trust and transparency.
- Ensuring that artificial intelligence does not violate civil liberties and the rule of law.
- Foreseeing and preventing unintended consequences of artificial intelligence use.
- Using artificial intelligence to expand equal opportunities, foster productivity and sustainable economic growth and enable access to new jobs, industries, education and innovation.

¹⁸ See also the OECD Principles on Artificial Intelligence.

- Promoting international collaboration and partnerships built on evidence-based approaches, analytical research and multi-stakeholder engagement.
 - Ensuring that research on and development of artificial intelligence are undertaken with the intention of assisting humanity in navigating the adoption and usage of these systems in an ethical and responsible manner.
58. Building a strong artificial intelligence governance framework will also require:
- **Human accountability.** Humans construct the algorithms on which the decisions informed by artificial intelligence systems are being made. Human accountability is therefore key to an ethical artificial intelligence approach.
 - **Regulatory compliance.** Artificial intelligence regulations help to protect user data and ensure responsible artificial intelligence use. Public and private organizations must adhere to data privacy requirements, accuracy standards and storage restrictions for safeguarding information. A central regulatory body for artificial intelligence, as proposed by the High-level Advisory Body on Artificial Intelligence, to ensure the responsible use of artificial intelligence, including in the public sector itself, could be considered.¹⁹
 - **Risk management and forecasting.** Artificial intelligence governance should include effective risk management strategies such as selecting appropriate training data sets, implementing cybersecurity measures and addressing potential biases or errors in artificial intelligence models. In addition, increased efforts are required to predict the possible disruptive effects that artificial intelligence might have in the future, paired with agile and forward-looking regulatory approaches.²⁰ Governments should also devise appropriate institutional mechanisms to anticipate risks and opportunities arising from artificial intelligence. The United Kingdom, for example, established the Regulatory Horizons Council, an independent expert committee that identifies the implications of technological innovation.
 - **Robust oversight and monitoring mechanisms.** Efficient oversight and monitoring mechanisms should be established to ensure the safe and responsible use of artificial intelligence. Upstream and downstream impact assessments of every algorithm should be conducted, including through ex ante social and ethical impact assessments.²¹
 - **Decision-making and explicability.** The decision-making capabilities of artificial intelligence systems should be fair and objective. To foster responsibility and trust, rationalization, or the capacity to comprehend the causes underlying artificial intelligence outputs, is crucial.
 - **Stakeholder involvement.** Governing artificial intelligence effectively will require stakeholder engagement in decision-making and oversight to ensure that artificial intelligence technologies are developed and used responsibly.
59. The future of artificial intelligence governance will rely on collaboration among all Member States and other stakeholders. Its success will depend on developing comprehensive artificial intelligence policies and regulations that protect the public while fostering innovation and bridging the gap that exists in the legal framework regarding artificial intelligence accountability, equity, transparency and integrity.

¹⁹ See also the 2023 interim report of the High-level Advisory Body on Artificial Intelligence.

²⁰ See for example OECD, “Recommendation of the Council for agile regulatory governance to harness innovation”, October 2021.

²¹ See also United Nations Educational, Scientific and Cultural Organization, “Recommendation on the ethics of artificial intelligence adopted on 23 November 2021”, 2022.

This will most likely also influence future regulation of other emerging technologies, such as biotechnology and neurotechnology.

VI. Conclusion and recommendations

60. Traditional and generative artificial intelligence are two distinct approaches in the artificial intelligence landscape. While the advantages of generative artificial intelligence lie in creativity, the handling of uncertainty, and novel applications, traditional artificial intelligence excels in efficiency, interpretability and specific task solving. Both approaches have their strengths and limitations, and their future in the artificial intelligence field holds tremendous potential for groundbreaking advancements and transformative applications.

61. Despite its multiple benefits, artificial intelligence is likely to pose a multitude of risks and challenges in the long run. On the cusp of a technological epoch, and to ensure that artificial intelligence contributes to the achievement of the 2030 Agenda, there is an urgent need for global cooperation, strategic foresight and an unwavering commitment to ensuring the equitable distribution of the benefits of artificial intelligence, while addressing its potential negative externalities. Only with a proactive, inclusive policymaking road map can the potential of this transformative technology be achieved.

62. Artificial intelligence has a special significance for the public sector itself, influencing the way in which policies are designed and government services delivered, as well as the transparency of decision-making. Going forward, further consideration needs to be given to the responsible and ethical use of artificial intelligence in the public sector.

63. Artificial intelligence governance is required, which should comprise a legal framework to ensure that artificial intelligence technologies are researched and developed with the goal of helping humanity to navigate the adoption and use of these systems in ethical and responsible ways. Artificial intelligence governance should aim to close the gap between accountability, transparency, ethics and integrity in technological advancement.

64. Ongoing efforts undertaken by governments, the United Nations system and other stakeholders, including the possible adoption of a global digital compact to be agreed at the Summit of the Future in 2024, should be continued and a global dialogue fostered to build the evidence base required to regulate artificial intelligence and ensure that it is aligned with shared global values, contributes to the achievement of the 2030 Agenda and leaves no one behind. The Group of 20 presidency of Brazil in 2024 could be another important opportunity to further the global agenda for a just digital transition and responsible use of artificial intelligence in favour of greater social inclusion and the reduction of inequality between and within countries. The year 2024 is a pivotal one for a fairer digital revolution with greater inclusion of vulnerable groups and developing countries.
