

GOVERNING WITH ARTIFICIAL INTELLIGENCE: ARE GOVERNMENTS READY?

OECD ARTIFICIAL
INTELLIGENCE PAPERS

June 2024 **No. 20**

This report was written by Barbara-Chiara Ubaldi, Deputy Head of Division and Head of the Digital Government Unit, and Ricardo Zapata, Policy Analyst, under the supervision of Carlos Santiso, Head of the OECD Innovative, Digital and Open Government (INDIGO) Division, Public Governance Directorate. Elsa Pilichowski, Director for Public Governance, and Gillian Dorner, Deputy Director for Public Governance, advised on the strategic approach and scope.

An initial draft of this report was enriched by the contributions and comments of the OECD Public Governance Committee in October 2023, as well as from the OECD Working Party of Senior Digital Government Officials (E-Leaders) in December 2023. The authors are thankful for the valuable comments from Miguel Amaral, Alana Baker, Charles Baubion, Erika Bozzay, Emma Cantera, Marco Daglio, Daniel Gerson, Pinar Guven, Andras Hlacs, Kenza Khachani, Eduardo Mangada Real de Usúa, Arturo Rivera, Alex Seeman, and the Trade Union Advisory Committee to the OECD. Further, the report was enhanced by the contributions from Karine Perset and Luis Aranda, among others, from the OECD Science, Technology and Innovation Directorate. Andrea Uhrhammer provided editorial support.

The paper was approved and declassified for publication by the OECD Public Governance Committee (PGC) on 11 June 2024 and prepared for publication by the OECD Secretariat.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

© OECD 2024

Cover image: © Kjpargeter/Shutterstock.com.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.

Governing with Artificial Intelligence: Are governments ready?

OECD countries are increasingly investing in better understanding the potential value of using AI to improve public governance. The use of AI by the public sector can increase productivity, responsiveness of public services, and strengthen the accountability of governments. However, governments must also mitigate potential risks, building an enabling environment for trustworthy AI. This policy paper outlines the key trends and policy challenges in the development, use, and deployment of AI in and by the public sector. First, it discusses the potential benefits and specific risks associated with AI use in the public sector. Second, it looks at how AI in the public sector can be used to improve productivity, responsiveness, and accountability. Third, it provides an overview of the key policy issues and presents examples of how countries are addressing them across the OECD.

Executive Summary

There is growing awareness that, if used strategically and responsibly, artificial intelligence (AI), including generative AI, has the potential to transform how governments function, design policies, and provide services. Governments have multiple roles in relation to AI, as enablers, funders, regulators, but also as users and in some cases as developers. While the global debate on AI has tended to focus on governments' role as regulators in shaping and responding to the application of AI, less attention has been paid to their responsibilities as users of AI. As governments seize the opportunities of AI for better governance and deploy solutions in a broad range of policy areas, they recognise the need to govern AI in the public sector to prevent misuse and mitigate risks.

In this context, OECD countries are increasingly investing in better understanding AI systems and using the opportunities they provide to transform the machinery of government. The responsible use of AI can improve the functioning of government administrations in several ways.

- First, the use of AI in the public sector can help governments increase **productivity** with more efficient internal operations and more effective public policies.
- Second, AI can help make the design and delivery of public policies and services more inclusive and **responsive** to the evolving needs of citizens and specific communities.
- Third, AI can strengthen the **accountability** of governments by enhancing their capacity for oversight and supporting independent oversight institutions.

This potential has by no means been fully explored and exploited. More evidence is needed on use cases to better understand how to successfully develop and deploy AI initiatives, learning from successes and failures. Despite the potential benefits of AI, there are also growing concerns about the risks of a fragmented and ungoverned deployment of AI in the public sector. Such risks include the amplification of bias, the lack of transparency in system design, and breaches in data privacy and security – all of which could lead to unfair and discriminatory outcomes with profound societal implications. The public sector has a special responsibility to deploy AI in a way that minimises harm and prioritises the well-being of individuals and communities, especially when deploying AI in sensitive policy domains such as law enforcement, immigration control, welfare benefits, and fraud prevention.

Governments are gradually working to establish an environment throughout the entire policy cycle to enable the safe, secure and trustworthy development, deployment and use of AI. These efforts comprehend defining strategic objectives, exploring new institutional arrangements, developing policy instruments (such as standards, codes, guidelines) and new regulatory frameworks, and attract the capacities needed to use AI effectively and efficiently in the public sector. Additionally, governments are increasingly prioritising monitoring implementation and impact to foster public trust and ensure long-term sustainability of current initiatives.

The OECD is supporting governments in their endeavour to pursue a safe, secure, and trustworthy use of AI in the public sector. The focus of the work is to guide governments in framing the relevant questions and identifying the most adequate enablers and guardrails that government need to consider when

choosing to apply AI to drive productivity, responsiveness, and accountability, as desired impacts. Fostering a common language will also help collecting evidence, fostering peer-learning and enhancing co-operation that will help governments be prepared to address common challenges.

This paper is part of a broader effort to understand the responsible use of AI in the public sector in a range of key government functions. As such, it seeks to ultimately contribute to the global need for further knowledge sharing, exchange of good practices, and structured policy dialogue to understand the implications and steer a responsible use of AI in the public sector. More and better indicators and evidence of the implementation and impact of AI on governments will help ensure it is used for optimal purposes; and multistakeholder engagement across policy sectors and beyond national borders will be needed to collectively explore policy options as new challenges and opportunities emerge.

1 Potential benefits and risks of artificial intelligence in the public sector

In addition to their role as regulators, governments are also users and, in some cases, developers of AI systems and applications¹ – progressively integrating them into their array of tools to deliver better policy outcomes (Ubaldi et al., 2019^[1]; Farrell et al., 2023^[2]; González, Ortiz and Sánchez Ávalos, 2020^[3]).² The expectation is that the use of AI can help governments improve in the following areas:

- **Productivity** in policymaking and service delivery, by:
 - Improving the **effectiveness** of policymaking, by using large amounts of data to gain more granular insights into user needs and identify patterns. This, in turn, would allow government to formulate more targeted policies and deliver better outcomes, by better targeting social expenditures, public investments and government services. For example, the municipality of Nijmegen in the **Netherlands** uses AI to count people in various locations of the city centre to monitor traffic and economic activity. The system outputs inform the design of more effective policies in areas such as road safety and entrepreneurial support.³
 - Increasing the **efficiency** of internal operations, by automating complex but repetitive administrative processes and procedures to support and facilitate the productive work of public officials, free up the time of skilled civil servants and ensure the reliability of the continuous delivery of public services. For example, the Queensland Government in **Australia** is using machine learning and computer vision to automatically map and classify land use features in satellite imagery, reducing the costs of mapping land use and improving the response to biosecurity and natural disaster events.
- **Responsiveness**, by improving governments' ability to anticipate societal trends and user needs to deliver proactive, personalised, and human-centred public services. For example, during the COVID-19 pandemic in **Norway**, the Labour and Welfare Administration used a conversational AI called Frida to help citizens access social benefits 24/7, resolving 80% of enquiries without requiring the intervention of a civil servant.⁴ This not only improved the quality of interactions by providing more timely and targeted assistance, which was crucial under the extreme circumstances, but also improved the civil servants' jobs quality by handling requests during non-standard hours and regular inquiries.
- **Accountability**, by using data analytics and machine learning techniques to detect fraud and risks to public sector integrity by identifying irregularities or suspicious patterns and raising red flags. For instance, AI is increasingly being applied to public procurement, public spending, and the provision of public grants, social benefits, and subsidies programmes. For example, Transport **Canada** has piloted the use of a risk-assessment algorithm to assess and identify potentially high-risk cargo before it is loaded onto inbound aircraft.⁵

At the same time, it is important to develop, deploy, and use AI in a safe, secure and trustworthy way in the public interest. Critical areas of concern include upholding human and civil rights, protecting personal privacy, securing algorithmic transparency, ensuring accountability, promoting “explainability”, and avoiding potential unfair and biased policy outcomes, among others. Some significant AI failures in the public sector have highlighted the need for governments to assess, test, and monitor AI’s impacts on the public. It is important to identify and manage risks for citizens; consider how AI systems may affect men, women or marginalized communities differently; ensure that the benefits of AI are distributed equitably; and mitigate potential harm.

In this context, governments are increasing their efforts to develop or procure trustworthy AI in line with the Principles of the OECD AI Recommendation (Box 1), that is, AI systems that respect human rights and privacy; are fair, transparent, explainable, robust, secure and safe; and where the actors involved in their development and use remain accountable (OECD, 2021^[4]). Governments are progressively taking steps to raise awareness and build capacity amongst civil servants, rethink governance mechanisms, upgrade regulatory frameworks, and strengthen central oversight and data governance. Some governments have even implemented outright bans of some types of AI applications in some jurisdictions (e.g., the use of facial recognition in San Francisco, along with various other municipalities in the **United States**⁶, and more recently, with the EU AI Act, as described in Box). Others have opted to provide guidance in the form of ethical frameworks and guidelines, standards, and codes of conduct (OECD, 2021^[5]). Sections 2 and 3 below further discuss the expected benefits and specific policy actions undertaken by countries to establish an enabling environment for safe, secure and trustworthy use of AI in the public sector.

Globally, the growing number of international standards highlight an emerging consensus for ensuring the responsible development and use of AI. Examples include the 2019 OECD AI Principles updated in 2024 (see Box 1), the Recommendation on the Ethics of Artificial Intelligence adopted in 2021 by the UNESCO⁷, the AI principles and “International Code of Conduct for Organizations Developing Advanced AI Systems” developed by the G7 as part of the Hiroshima AI Process⁸, the EU AI Act⁹, the Council of Europe’s Framework Convention on Artificial Intelligence¹⁰, and the 2024 UN Resolution on the promotion of safe, secure and trustworthy artificial intelligence systems for sustainable development.¹¹

However, these initiatives tend to indirectly address the public sector, underscoring the need for more specific guidance. More recently, the G7 “Ministerial Declaration on Industry, Technology, and Digital” of March 2024 recognised the need to ensure that public sectors are equipped to deal with AI systems.¹² In Ibero-American countries, a more targeted standard has been developed through the regional Charter on Artificial Intelligence in Public Administration (Box 2). To address the need for more targeted international guidance on AI for the public sector, Section 4 presents a preliminary framework for the implementation of trustworthy AI.

Box 1. OECD AI Principles

The OECD Principles on Artificial Intelligence support AI that is innovative and trustworthy, and which respects human rights and democratic values. OECD Member countries adopted the Principles in May 2019 as part of the OECD Council Recommendation on Artificial Intelligence and updated them in May 2024 [OECD/LEGAL/0449]. They complement existing OECD standards in areas such as privacy, digital security risk management and responsible business conduct. The Recommendation identifies five values-based principles for the responsible stewardship of trustworthy AI and provides five recommendations to governments.

Values-based principles	Recommendations to governments
Inclusive growth, sustainable development and well-being	- Investing in AI R&D - Fostering an inclusive AI-enabling ecosystem

• Human rights and democratic values, including fairness and privacy • Transparency and explainability • Robustness, security and safety • Accountability	• Shaping an enabling governance and policy environment for AI • Building human capacity and preparing for labour market transition • International co-operation for trustworthy AI
--	---

Source: <https://oecd.ai> and (OECD/CAF, 2022^[6]).

Box 2. Ibero-American Charter on Artificial Intelligence in the Public Administration

The Charter was issued by the Latin American Centre for Development Administration (CLAD), an intergovernmental organization comprised of 24 Latin American countries, Spain, and Portugal. It was formulated to foster a common framework for the development and application of AI within public administrations in the Ibero-American region. The Charter sets forth guiding principles and outlines key dimensions and directions for the holistic and systematic adoption and use of AI across all state bodies and institutions.

Source: <https://clad.org/declaraciones-consensos/carta-iberoamericana-de-inteligencia-artificial-en-la-administracion-publica/>

2 Delivering impact by leveraging artificial intelligence in the public sector

This section considers how countries are exploring and exploiting the potential of AI in the public sector, based on the analysis of concrete use cases. To gain a deeper understanding of how the use of AI in the public sector is expected to produce benefits and help delivering impact, Table 1 presents a preliminary taxonomy¹³ applied to analyse the use cases. This taxonomy outlines the specific tasks performed by AI systems, in line with the *OECD Framework for the Classification of AI Systems* (OECD, 2022^[7]). Second, the taxonomy distinguishes four key functions of the public sector where these tasks can be performed:

- Efficiency - Improve the **internal operations** of public administrations (where AI is most commonly used).
- Effectiveness – Improve **policy making**, which is understood as the decision-making process to plan, implement, and (where relevant) alter public policies and programmes (OECD, 2020^[8]).
- Responsiveness -Improve **service delivery**, which is understood both in terms of the provision of public services and their design.
- Accountability - Enhance **oversight and risk detection**, both within government agencies in high-risk areas and by external oversight agencies (OECD, 2024^[9]).

Finally, the taxonomy looks at the impact of these uses cases for the public sector in terms of responsiveness, productivity (efficiency and effectiveness) and accountability.

Table 1. Understanding the use of AI in the public sector

Tasks	Function	Impact
• Recognition. • Event detection. • Forecasting. • Personalisation. • Interaction support. • Goal-driven optimization. • Reasoning with knowledge structures. • Content generation	Internal operations	Productivity (efficiency and effectiveness)
	Policy making	Responsiveness
	Service delivery	
	Internal and external oversight	Accountability

Source: authors' elaboration, using "AI System Tasks" taxonomies from (OECD, 2022^[7]).

The taxonomy was employed to analyse selected examples, chosen from a pool of 71 use cases¹⁴ across 31 Member and accession countries. These were collected in 2023 as part of the *OECD Survey on Digital*

Government 2.0 and the *Call for Innovations in Government* of the OECD's Observatory of Public Sector Innovation (OPSI).

Tasks

In most of the cases reviewed, AI is used to perform specific tasks, such as recognition (i.e., identifying and categorising data), forecasting (i.e., predicting future outcomes), and interaction support (e.g., chatbots and virtual assistants). Box 2.1 provides detailed definitions and examples of these tasks.

For instance, in **Luxembourg**, the Information and Press Service has employed facial recognition technology to identify public figures (i.e. politicians) within an extensive media library of political photos. This initiative aims to enhance the metadata associated with the media archive, ultimately improving public communication: with properly tagged photos, media representatives and journalists can swiftly access visual materials for their reporting, thereby facilitating more accurate and timely news coverage. Another example is the use of AI by civitech initiatives deployed by governments for event detection and recognition tasks. This is the case of CitizenLab in **Belgium**, which is used by 400+ local governments to enhance their understanding of and responsiveness to citizens' feedback. It enables civil servants to better group and categorize citizen input collected through online platforms, streamlining the processing of thousands of contributions, and improving the ability to address community needs more effectively. The system allows to better analyse public consultations across different policy areas (e.g., environment, urban planning, local government, infrastructure).

Box 2.1. Tasks of AI Systems

The task of an AI system is the most basic job that the AI is programmed to perform. They can be categorised into seven main types of tasks that most AI systems can do and which can be applied in the public sector:

Task	What it does	Examples
Recognition	Identifies and categorises data (e.g. image, video, audio and text) into specific classifications. Output is often one label, e.g. “this is a cat”.	Image & object detection; facial recognition. Audio, sound, handwriting and text recognition; gesture detection.
Event detection	Connects data points to detect patterns as well as outliers or anomalies.	Fraud and risk detection, flagging human mistakes, intelligent monitoring.
Forecasting	Uses past and existing behaviours to predict future outcomes, generally to help make decisions. Contains a clear temporal dimension.	Assisted search, predicting future values for data, predicting failure, predicting population behaviour, identifying and selecting best fit, identifying matches in data, optimising activities, intelligent navigation.
Personalisation	Develops a profile of an individual and then learns and adapts to that individual over time. The output is usually a ranking, e.g. a search engine ranking.	Recommender systems based on search and browsing (Netflix, Amazon), personalised fitness, wellness, finance.
Interaction support	Interprets and creates content to power conversational and other interactions between machines and humans (e.g. involving voice, text, images). Can be realtime or not.	Chatbots, voice assistants, sentiments model and intent analysis, back-office process automation (rule-based routing, moving information between systems)
Goal-driven optimisation	Gives systems a goal and the ability to find the optimal solution to a problem, which can be by learning through trial and error. It assumes a cost function is given.	Game playing, resource/logistics optimisation, iterative problem-solving, bidding and advertising, real-time auctions, scenario simulation.
Reasoning with knowledge structures	Infers new outcomes that are possible, even if they are not present in existing data, through modelling and simulation.	Expert systems, legal argumentation, recruitment systems, diagnosis, planning.
Content generation	Refers to the autonomous production of text, images, audio, or video from input data, often leveraging natural language processing and deep learning techniques.	Writing articles, generating product descriptions, creating artwork, producing synthetic speech, composing music, generating deepfake videos.

Source: adapted from (OECD, 2022^[7])

Government functions

Internal operations

In most of the use cases reviewed, AI is mainly used to increase efficiency of internal operations in the public sector. This is consistent with the evidence collected through the *2023 OECD Digital Government Index*, indicating that about 70% of participating countries have used AI to enhance internal operations (OECD, 2024^[10]).

For example, **France** is currently experimenting with a generative AI¹⁵ tool called “Albert” to streamline the daily tasks of France advisors on public services. This tool aids in accessing information, related sources, frequently asked questions, and practical links, enabling advisors to provide customised assistance to the users they assist¹⁶. The tool is characterised by its free and open-source approach. Furthermore, France has initiated a pilot project using AI to assist managers in recognising the skills required for assessing public servants' missions. **Canada** is using robotic process automation to automate tedious tasks such as transferring information between systems, streamlining internal operations and increasing efficiencies of officers' workflows. These tools are used in multiple federal departments and principally support interaction. In **Sweden**, the Companies Registration Office developed an AI model that sorts

approximately 60% of incoming emails¹⁷. The model reads their content, detects specific key phrases, and forwards it to the right recipient within the Office. In the case that an email does not contain one of the predefined key phrases, it reviews its entire content and makes an assessment based on employees' previous behaviours. In **Italy**, the *Corte dei Conti* (Court of Auditors) uses a custom-AI model called GiusBERTo to automatically deidentify and anonymise court decisions without sacrificing any important information, a process previously done manually. This solution helps to balance the public's right to access information with the need to protect the privacy of citizens. The anonymised documents are then subject to human review to ensure their accuracy and completeness (Datta et al., 2023^[11]).

As countries acknowledge the potential that AI can bring to enhance public sector productivity, they are adapting their internal frameworks to facilitate its systematic use. For instance, the Infrastructure and Projects Authority (IPA) of the **United Kingdom** has released a framework to use AI for enhancing productivity in the delivery of public projects, while also promoting responsible experimentation with AI to address commonly major challenges in public projects¹⁸. The framework recognises the necessity of enhancing the enabling environment, including the upskilling of project delivery professionals in AI and data analytics (see Section 3 below).

Policy making

Many use cases aim to improve effectiveness in policymaking. This is seen particularly in tax administration (e.g., streamlining tax collection processes and preventing tax fraud), budget management (e.g., analysing spending and revenue data to develop and manage budget more effectively, and anticipating budget allocations), public health, welfare benefits, and social services. However, according to the *2023 OECD Digital Government Index*, only 30% of countries have used AI for policymaking, a significantly lower adoption rate compared to its application for internal operations (OECD, 2024^[10]). One example from the public health sector is the AI convergence system developed by **Korea's** Disease Control and Prevention Agency to address situations of emerging infectious diseases. The system performs forecasting tasks by analysing medical data, quarantine data, and spatial data to develop policy responses to infectious diseases.

Although less prominent across the use cases analysed, governments are also using AI to improve policymaking in other sectoral areas, such as transport, environment, security, urban planning, employment, justice, infrastructure, education, foreign affairs, among others. For example, in **Colombia**, a project developed by the Organisation of Ibero-American States with the National Land Agency combined AI, satellite imaging, and online monitoring platforms to generate detailed geo-cadastral information within a 200,000-hectare polygon in a remote rural area, allowing the recognition and registry of nearly 1,900 properties.¹⁹ In **France**, the Paris-Saclay agglomeration of municipalities is using AI to simulate different energy management scenarios through a digital twin of their territory, allowing officials to more effectively evaluate the environmental and financial impacts of projects and improve long-term planning capabilities²⁰.

Service design and delivery

Governments are also using AI to increase responsiveness in the design and delivery of public services. Similar to the uptake of AI use for internal operations, 67% of OECD countries are using AI to improve the design and delivery of public services, according to the *2023 OECD Digital Government Index* (OECD, 2024^[10]). For example, **Finland** is using the AuroraAI programme to identify public services that are overly cumbersome for the user²¹. It uses AI to simulate potential service paths and proactively offer citizens services based on life events (e.g. marriage, beginning university, retirement). This system is being used across many policy areas and performs mainly tasks for recognition and personalisation²².

At the sector level, the **Austrian** Digitalisation and E-Government Directorate of the Federal Ministry of Finance developed Mona, a conversational chatbot to provide information to entrepreneurs about

business-related services and help them navigate the most relevant web content, increasing service quality and relieving civil servants from excessive workload. The system improves responsiveness in public services and performs principally interaction support tasks. In **Türkiye**, the Ministry of Health integrated an AI-powered application called "Neyim Var?"²³ (or *What's wrong with me?*, in English) within "e-Nabız", an e-health platform offering access to various health services. The AI application provides personalised guidance to citizens, including healthy nutrition recommendations. Public employment services or labour agencies are also using AI systems to improve service delivery. For example, the **France's** Pôle Emploi, uses a tool called "Automatic CV Analysis" (Analyse Automatique de CV – AACV) which helps jobseekers from the moment they create a profile. In **Belgium**, the Public Employment Service of Flanders' (VDAB) Competency-Seeker platform helps both jobseekers and employers enrich and refine the skills profiles they have and are looking for (Broecke, 2023^[12]).

Government oversight

A variety of use cases show that AI can help increase transparency and accountability of government activities, when tested appropriately and used responsively (OECD, 2022^[13]). Tax and procurement agencies, as well as audit institutions have been early adopters of these technologies, in areas ranging from fraud detection and corruption control to risk management.

For example, **Spain's** Comptroller General has used AI to identify high-risk instances of potential fraud in grant and subsidies programmes (OECD, 2021^[14]). In **Estonia**, the Tax and Customs Board (MTA) has been testing AI to identify incorrectly submitted VAT refund claims and to identify companies or persons in need for inspection²⁴. Such use cases perform mainly event detection tasks. An interesting example of sectorial oversight is the data science application, SELFIM, built by the **French** Ministry of Interior, that automatically detects attempts to defraud the vehicle registration certificate (VRC). Other central oversight agencies and independent watchdogs have been using AI to better identify suspicious patterns and raise red flags, in particular with government procurement. In **Brazil**, for example, the General Controllershship of the Union created the Analyzer of Biddings, Contracts and Notices (Alice) to deliver preventive and timely action in relation to public procurement. The system automatically collects information on the ongoing processes on the main public procurement platforms of the Federal Government on a daily basis, assesses a set of risks and issues, alerts to direct the attention of the auditors and managers involved to situations that do not conform to the expected standards (Oliveira, Monteiro da Rocha and Scatolino de Rezende, 2022^[15]).

3 Policy issues and actions for a responsible use of artificial intelligence in the public sector

Governments are working to establish an environment that enables the safe, secure and trustworthy development, deployment and use of AI to maximise benefits and address associated challenges. This section explores the main policy challenges emerging for the public sector and the actions developed by OECD countries to address them.

Defining strategic objectives

Securing political and public support to scale-up the deployment of AI in the public sector remains a challenge in most countries. To gain and maintain the support needed, governments are developing strategies for AI adoption in the public sector. The intent is to steer the coherent development, deployment, and use of AI across public sector institutions in line with overarching values and objectives. Governments are also exploring new institutional arrangements or adjusting the mandates of existing institutions, such as those responsible for digital government, data governance, public sector innovation, government transparency or oversight, to integrate AI in their mandates and functions. Furthermore, Parliaments are increasingly turning their attention to the responsible use of AI in and by public sectors and are creating pressure on governments to provide transparent and strategic steering, e.g. special inquiry commissions and/or committees have been established in Parliaments in **Australia** and in the **United Kingdom**.

Several actions are expected to advance a more coherent and trustworthy development, deployment and use of AI in and by the public sector. These include:

- providing a **whole-of-government approach** through strategic guidance, clear objectives and oversight to steer the integration and procurement of AI across sectors, ministries, and agencies in accordance to expected values and guiding principles.
- embedding **participatory mechanisms** within institutional arrangements to increase the engagement of and oversight by citizens and other relevant stakeholders in the design, development, and deployment of AI in the public sector (Wong et al., 2022^[16]).
- establishing robust **data governance** for the whole public sector to support access to and sharing of high-quality government data, often integrating data governance strategies with AI strategies; acknowledging that trustworthy use of AI relies on high quality data.
- ensuring the **reliability** of AI systems used in the public sector. This entails developing or using the right metrics and tools²⁵ to systematically scrutinise AI systems and models used in the public sector to make sure they deliver inclusive and fair results. This includes also having safeguards in place to prevent or rectify biases (e.g., gender, racial, etc.) and other problematic outputs.

- enhancing **accountability** in the deployment of AI within the public sector through greater transparency, explainability, and traceability, for instance through the establishments of algorithmic accountability standards and open registries of public algorithms²⁶.

To better respond to these expectations, governments are exploring new institutional arrangements. These arrangements aim to ensure the necessary steering from the centre of government for a coherent and accountable deployment and use of AI across the public sector, in line with overarching democratic values and the pursuit of the public interest. A number of several European countries are expanding the mandates of existing ministries or agencies, such as **Norway's** Ministry of Digitalisation and Public Governance²⁷ or Spain's Secretary of State for Digitalisation and Artificial Intelligence within the Ministry of Digital Transformation and Public Management. A similar trend is emerging in Latin America. A review on the use of AI in the public sector indicates that at least half of the countries in the region have identified a specific government organisation to drive AI efforts (OECD/CAF, 2022^[6]).

Additionally, new institutions are emerging, such as the European AI Office²⁸, the European Artificial Intelligence Board, the European Centre for Algorithmic Transparency (ECAT)²⁹ or **Spain's** AI Supervision Agency³⁰. These institutions are tasked with supervising the implementation and enforcement of regulations related to AI, in alignment with the new EU AI Act requirements (see Box 4). Complementary arrangements introducing participatory mechanisms are also being explored, such as the citizen assembly on AI launched in the EU by the Belgium Presidency³¹. However, the mandate of these institutions in relation to the use of AI in the public sector is not always clearly defined. Other countries are establishing new roles to ensure leadership and oversight for the use of AI in government, often adopting decentralised approaches. The **United States** requires federal agencies to designate Chief AI Officers, responsible to coordinate the use of AI across their agencies, and to establish AI Governance Boards, chaired by the Deputy Secretary or equivalent, to coordinate and govern the use of AI across the agency³².

There are also attempts to leverage existing arrangements to set up new coordination and regulatory mechanisms. **Australia** for example created an AI in Government Taskforce under the joint leadership of the Digital Transformation Agency and the Department of Industry, Science and Resources³³. The taskforce has been mandated to develop guidelines and a governance approach on how to best enable the safe, ethical, and responsible use of AI in public service, including to improve risk management, skills and capability, technical use, and preparedness. In the **European Union**, some member states have opted to create a unit in charge of algorithms inside their national Data Protection Authorities to be the AI regulatory bodies (OECD, 2023^[17]).

While many countries have similar efforts underway, more analysis is needed to better understand which institutional structures and organisational arrangements are more effective in steering, managing, and overseeing the deployment of AI in a given administrative context. This may imply the establishment of new institutions or broadening of the scope of the mandate of existing ones to drive the effective implementation of AI systems, thereby enhancing their accountability and ultimately ensuring long-term viability. When looking into options the overall institutional context of a country should be considered to ensure flexibility to reinforce coordination and oversight across policy sectors and levels of government.

In addition to new institutional arrangements, almost all OECD countries are developing or updating their strategies, agendas or plans for AI that include specific objectives or actions for its use in the public sector (OECD, 2024^[10]). The most common objectives of these policy documents, that are specifically designed to foster a coherent deployment and use of AI in the public sector, include improving digital public infrastructure, strengthening data governance, promoting AI adoption, and enhancing AI governance. Developing cross-border ecosystems and strengthening civil servants' capabilities are less developed aspects. When looking at the enablers to help drive progress in AI deployment and greater accountability, most of the reviewed strategies define specific actions and set goals, while about half of the strategies establish funding mechanisms, develop monitoring instruments, and define responsible actors. A minority sets implementation timeframes for specific actions.

However, AI strategies specifically targeted to the public sector are still very limited. Most of the existing strategies cover a broader scope and actions to catalyse economic development, transform the labour market, or strengthen AI skills and talents across society. Among those efforts more targeted to the public sector, in **France**, the *Conseil d'État* published in 2022 a report that supports the development of a dedicated strategy for trustful AI in the public sector to improve public performance and the quality of public services³⁴. As a more targeted instrument, the country's General Directorate of Administration and Civil Service (DGAFP, by its acronym in French) is currently developing a strategy for the use of AI for human resources management (HRM) in the public sector³⁵. The strategy seeks to support experimentation with AI actions to enhance empowerment and service quality, identify risks related to AI adoption in public HRM, assess current and future AI skill needs, among others. Outside the OECD, **Uruguay's** AI strategy³⁶ has been highlighted due to its dedicated focus on the public sector (OECD/CAF, 2022^[6]).

Designing policies

Risks associated with the use of AI (see Section 1) and unsuccessful use cases have sparked the development of dedicated policies, standards, codes, and guidelines, and in some cases new regulatory frameworks. These initiatives aim to more effectively prevent and address failures or to guide the ethical and responsible development of AI systems in the public sector. The approaches from countries in driving good use include soft guardrails, as well as new regulations and legislation. However, enforcement is still a challenge, either concerning the correct application of these policies or the availability of the necessary capabilities to drive accountability.

To address these concerns, most countries have set a variety of soft guardrails like ethical frameworks, standards, guidelines, or transparency tools to steer the responsible deployment of AI in the public sector. For example, **Australia** developed policy tools, such as the "Artificial Intelligence Ethics Framework" and the "Automated decision-making Better Practice Guide", aimed at providing direction to government officials for the ethical use and management of algorithms. In **Colombia**, the Ethical Framework for Artificial Intelligence³⁷ offers a series of principles, along with a methodology for their implementation, that should be considered in the design, development, and implementation of AI systems. **Canada's** Guide on the use of generative AI³⁸ advises public sector organisations to be aware of amplification of biases that might be dominant in the training of data and requires to mitigate them from the planning and design stage. **Japan** developed the "AI Guidelines for Business" in April 2024, outlining critical factors to consider when developing, providing, or using AI³⁹. These Guidelines are intended for all AI business actors, including public institutions such as governments and municipalities. They incorporate the international guiding principles and international code of conduct for advanced AI systems, as compiled in the Hiroshima AI Process. Other tools include **France's** guide for public algorithms' transparency⁴⁰, **Ireland's** Interim Guidelines for Use of AI in the Public Service⁴¹, the **Netherlands'** guide on governance for a responsible application of AI⁴², and the **United Kingdom's** Algorithmic Transparency Recording Standard⁴³. At the sub-national level, some cities have been adopting standards on digital rights, such as the Cities Coalition for Digital Rights⁴⁴, representing over 50 cities worldwide. These standards are being put into practice through projects like open registries of public algorithms⁴⁵, a Global Observatory of Urban AI⁴⁶ or a Digital Rights Governance Project⁴⁷.

Various countries are updating their regulatory frameworks to support a safe, secure and trustworthy AI use in the public sector. For example, **Canada's** federal government issued the Treasury Board Directive on Automated Decision-Making⁴⁸, a mandatory policy instrument to ensure that automation in administrative decision-making in the federal public service is compatible with administrative legal principles such as transparency, accountability, legality, and procedural fairness. The Directive also requests that the Gender-based Analysis Plus⁴⁹ analytical tool is used during the development or modification of an AI system to understand its impact on different population groups. Similar efforts can be observed elsewhere, such as the United States' 2024 Executive Order on Advancing Governance,

Innovation, and Risk Management for Agency Use of Artificial Intelligence⁵⁰ setting guidelines for the responsible use of AI across the federal government. In the **European Union**, the adopted AI Act⁵¹ establishes a novel regulatory framework, also impacting the use of AI systems in the public sector (see Box).

Box 2. The EU AI Act and its implications for the public sector

The EU AI Act is a European regulation on AI agreed in negotiations with member states in December 2023 and approved by the European Parliament in March 2024. The regulation establishes obligations for AI based on its potential risks and level of impact. Furthermore, the Act fosters a reformed institutional architecture both within individual countries and at the European level. The Act identifies different levels of risks which are relevant for governments' use of AI.

Risk levels and obligations

The AI Act defines four risk levels:

- **Unacceptable risk:** AI uses under this category are prohibited by the AI Act. Examples include predictive policing, 'real-time' remote biometric identification (including facial recognition) in publicly accessible spaces for law enforcement, social scoring, or assessing the risk of an individual committing criminal offenses. Law enforcement and justice are among the public sector policy areas most concerned by this category, although some exceptions apply, such as use cases concerned with national security and those remaining subject to judicial oversight.
- **High-risk** - AI uses under this category are allowed but regulated due to their significant potential harm to health, safety, fundamental rights, environment, democracy, and the rule of law. Due to its potential impact on these aspects, most public sector uses of AI might fall under this category. Examples include systems used to influence the outcome of elections and voter behaviour, automated processing of personal data to assess various aspects of a person's life, assessing eligibility to benefits and services, and safety components used in the management and operation of critical infrastructure. Obligations include establishing a risk management system, conducting data governance, having in place technical documentation to demonstrate compliance, mandatory fundamental rights impact assessment, among others.
- **Limited risk** – These systems might include chatbots, deep fakes, emotion recognition systems, among others, and have transparency obligations where developers and deployers must ensure that end-users are aware that they are interacting with AI.
- **Minimal risk** – These systems are unregulated, but a code of conduct is suggested. Examples include video games and spam filters.

Governance Framework

The Act also introduces a restructured governance framework at both national and European levels. Each member state must designate one or more **National Competent Authorities** to supervise the Act's enforcement. At the European level, the **European Artificial Intelligence Board** will gather official points of contact of each Member country to ensure uniform application across member states. It will be complemented by an **advisory forum**, representing a balanced selection of stakeholders, and a new European **AI Office**, to be established within the Commission, which will be supported by a **Scientific Panel of Independent Experts**.

Source: (Future of Life Institute, 2024^[18]); <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>, <https://www.adalovelaceinstitute.org/wp-content/uploads/2022/04/Expert-explainer-The-EU-AI-Act-11-April-2022.pdf>, https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683

Supporting implementation

Public sector organisations are being encouraged to test AI systems, but numerous implementation challenges persist, as AI remains largely experimental and there is not yet a robust track record of systematic use in the public sector, nor on the resources (human and financial) required to scale initial efforts (OECD, 2024^[10]). Countries have been deploying diverse initiatives to enhance their capacity to use AI effectively and efficiently in the public sector, including the following:

- Developing core digital public infrastructure to support AI use and development and help increase traction and adoption. For instance, **Estonia** is developing three sets of reusable core components for service development: first, data and open data catalogues and portals; second, data processing tools (e.g. consent service⁵², data tracker⁵³, and anonymisation); and third, basic AI tools to be re-used across various other AI applications (e.g. virtual assistant⁵⁴, speech synthesis, translation tools, text analysis toolkit, among others). Findings from the *2023 OECD OURdata Index* also reveal that OECD countries have improved the quality of open government data, which can be instrumental for the development of AI models (OECD, 2023^[19]).
- Establishing and maintaining a digitally capable workforce, which requires the public sector to become better at attracting and retaining relevant digital talent, in addition to offering skills development programmes for civil servants to increase their awareness on potential risks and increase their capacity to take informed decisions on the procurement, development, deployment and use of AI. For example, **Finland's** Elements of AI⁵⁵ is a free and open course to help both citizens and public servants gain a solid understanding of AI. In **Estonia**, the skills development programme includes a data expert network with 500+ participants, AI meetups, experimentation events (e.g., hackatons, competitions), and public e-courses on AI. In **Ireland** an AI upskilling programme for civil servants has been in place since 2021 as part of the public service transformation strategy, 'Better Public Services'. Civil and public servants interested in learning how AI could be used to transform public service delivery have access to short courses. Among its offerings is an online course that guides participants in designing potential AI projects that could benefit from strategic investment⁵⁶.
- Securing a whole-of-government approach to digital government investments to maximise the impact of AI systems and support their successful implementation. By developing a coordinated approach when investing in the digital transformation of the public sector, governments can promote a coherent and strategic adoption of AI in the public sector, mitigating potential risks and securing value-for-money. A structured approach to digital government investments can help governments deploy AI solutions coherently across government, for example by leveraging value proposition and funding to secure standards compliance in terms of algorithmic transparency, trustworthy and responsible use of data. Governments can also adopt dedicated funding mechanisms to promote AI adoption across the administration aligning individual investments with strategic goals. For example, **Estonia** has successfully integrated AI solutions into its public service infrastructure, enhancing responsiveness and reliability for its citizens. It has also created flexible and dedicated funding opportunities to support the uptake of AI in the public sector, including through structural funds, joint procurements, and new upcoming funding measures⁵⁷.
- Enhancing the capacity for effective procurement of AI by government entities can stimulate innovation and shape the development of AI solutions aligned with government policy and societal

values, including through partnerships with GovTech startups (Farrell et al., 2023^[2]). Procurement standards such as IEEE P3119⁵⁸ are being developed to help public organisations strengthen their requirements for AI procurement in the public sector. Governments are also standardising contractual clauses to support public organisations wishing to procure AI systems developed by external suppliers. For example, in 2023, the European Commission published the EU model contractual AI clauses to provide a clear and consistent framework for AI procurement, ensuring that all parties have a common understanding of the terms and conditions⁵⁹. These represent a proposal for standard contractual clauses for the procurement of AI, developed by a community of public buyers. There are two versions of these clauses: one for high-risk AI systems and another for non-high-risk systems. Public procurement rules can also be leveraged to embed ethical standards in how governments purchase AI solutions. For instance, the Guidelines for AI Procurement in the **United Kingdom** includes a set of principles on how to buy AI technology, as well as insights on tackling challenges that may arise during procurement⁶⁰. **Canada** opted for another approach to support the ethical purchase and deployment of AI through the establishment of a list of pre-qualified suppliers who can provide responsible and effective AI services, solutions and products, streamlining procurement processes of public organisations that need to comply with the country's Directive on Automated Decision-Making⁶¹.

- Leveraging partnerships with actors from outside of the public sector (e.g., scientific communities, or academic institutions) to gain external expertise for the development of AI policies or initiatives. For instance, the **United Kingdom** does this through the Public Policy Programme⁶² at The Alan Turing Institute. In the **United States**, the TechFAR Handbook⁶³ highlights flexibilities in the country's 2,000+ page procurement regulations, allowing agencies to work with start-ups and conduct iterative, user-driven service development.

Securing monitoring and oversight

Finally, monitoring the effectiveness of policies for AI in the public sector and assessing their implementation and impact is crucial to foster trust, and ensure their long-term sustainability. Currently, setting-up transparency, monitoring and oversight mechanisms on AI in the public sector remains a challenge in most countries. Some countries are exploring different mechanisms and institutions, for:

- Monitoring the use and impact of AI in the public sector to understand if the promise of better delivery of policies and services, improved decision-making processes, and greater productivity is being fulfilled. At a global level, the OECD AI Incidents Monitor (AIM) is being used to keep track of AI incidents and hazards, including in the public sector. Its goal is to assist policymakers, AI practitioners, and other stakeholders gaining valuable insights into the risks and harms of AI systems.⁶⁴ Countries are also developing their own monitoring mechanisms to assess AI impact. For instance, in **Türkiye**, the Digital Transformation Office of the Presidency of the Republic conducts the "AI Risk Management Recommendation" and "Trustworthy AI Seal" studies to closely monitor the use of AI for public benefit. As a baseline for impact measurement, many governments are also developing standards for algorithm accountability in the public sector, such as algorithm registers. The purpose is to ensure the transparency of the data used to train AI systems, of the objectives they are used for, or of how potential impacts or risks on individuals or society are being assessed. Registries also provide citizens a way to evaluate or question governments' application of AI, but there are few cases at the central/federal level. For instance, **Canada's** requires the publication of completed Algorithmic Impact Assessments⁶⁵ and the Netherlands' Algorithm Register⁶⁶ has similar objectives. Measuring the actual impact of AI projects remains an area of opportunity across countries, with the need to go beyond the baseline provided by algorithmic transparency instruments. As a response to this challenge at the **European Union** level, a report by the Joint Research Centre recommended adopting value-oriented AI impact assessment co-

created frameworks that prioritise *sustainability* over mere return-on-investment, “including optimisation of savings in terms of human, financial and environmental costs, increased efficiency, enhanced effectiveness in service delivery, service quality and trustworthiness, improved relevance and contextualisation” (Manzoni et al., 2022^[20]).

- Securing oversight of the proper development and use of AI within the public sector by non-executive branches of government (e.g., judiciary and parliament as mentioned above) and accountability institutions (e.g., access to information agencies, data protection agencies, ombudspersons, audit offices), which is progressively gaining traction⁶⁷. For example, **Norway’s** Office of the Auditor General (OAG) is auditing the use of AI in the central government since 2023 as part of its pipeline of new performance audits⁶⁸. Additionally, in its Strategic Plan 2018-2024, the OAG also envisions using AI for service delivery, highlighting that “problem solving will become more automated, and the use of [AI] will gradually take over tasks in both the public administration and the OAG”⁶⁹. In the **United Kingdom**, the National Audit Office (NAO) reported in March 2024 that 70% of the government bodies that were surveyed were piloting or planning the use of AI to support operational decision-making and improve internal processes⁷⁰. The report highlighted the potential risks to value for money that may arise in the absence of a clear identification of the institution responsible for the implementation of the AI adoption strategy for the public sector. This may hinder the wide-scale benefits achievable through tackling jointly common issues of ageing IT infrastructure, risks, skills gaps, and data issues. Around the world, countries are exploring new accountability structures, such as the **Spanish** new independent Artificial Intelligence Supervision Agency⁷¹ also mentioned above, or instruments for auditors, such as the frameworks for assessing whether algorithms meet quality criteria by the **United States’** Government Accountability Office (GAO)⁷² and Court of Auditors (NCA) in the **Netherlands**⁷³.
- Enforcing regulations across public sector institutions, with varying approaches, can also help promote the development of risk-based approaches. While the EU AI Act foresees its enforcement under single national regulators, in other instances the responsibility for regulatory delivery and enforcement may be shared between several institutions. The **United Kingdom** for example is laying the groundwork for its future model of AI regulation along with its new Data Protection and Digital Information Bill. The plan is to give responsibility to several institutions the UK communications regulator (OFCOM), the Competition and Markets Authority (CMA), the Information Commissioner’s Office (ICO), the Financial Conduct Authority (FCA) and the Medicines and Healthcare products Regulatory Agency (MHRA) are on the list, with some that can expect their competences and powers to be updated.
- Strengthening transparency to enhance accountability, based on the understanding that functional transparency mechanisms are fundamental to challenge or seek redress for an outcome of an AI system (OECD, 2023^[21]). Current data protection and access to information legislation and institutions are already present in 134 countries worldwide, including 37 OECD Members (OECD, 2022^[22]). This can be used in some instances to strengthen AI transparency, in addition to the initiatives mentioned above focused on the monitoring of impact and use, e.g. such as the provisions for ex ante and ex post (social and ethical) impact assessments seen as important safeguards. However, more specific requirements and involvement from non-executive oversight institutions might also be needed to increase proactive disclosure of AI algorithms by the public sector and provide means of recourse in case of misuse. For instance, **Chile’s** Transparency Council is developing a General Instruction on Algorithmic Transparency that will mandate more than a thousand public agencies to disclose the algorithms they employ in providing services to the population, allowing citizens to understand whether the services they receive are driven by algorithmic models and understand the underlying decision-making logic⁷⁴. Furthermore, to enhance transparency in the use of AI, several countries are creating inventories of use cases. These include the **United States** where public sector institutions are required to annually release

public inventories of AI use cases, and to highlight the ones that impact rights or safety and indicate how the agency is addressing the relevant risks⁷⁵.

These policy issues and the diversity of approaches highlight the need for further sharing of knowledge and exchange of good practices in terms of countries' policies and practices to ensure the responsible deployment of AI in the public sector. This can foster a structured policy dialogue amongst OECD member and partner countries to better understand the implications and steer a responsible use of AI in the public sector. The following section presents a preliminary framework conceived for this purpose.

4 Towards a framework for a trustworthy use of AI in the public sector

The OECD is developing a framework to support governments in the responsible use of AI in the public sector. This framework outlines the essential enablers and guardrails that governments can consider to address existing constraints for a trustworthy use of AI. Moreover, it emphasises the engagement of diverse stakeholders in both the design and implementation phases of AI initiatives in the public sector to increase productivity, responsiveness, and accountability, as desired impacts (Figure 1).

The proposed framework is based on the evidence collected by the Secretariat and is intended to foster dialogue on how governments can frame the development and deployment of AI in the public sector in a coherent, safe, secure and trustworthy manner. Given the rapid development and evolving nature of AI, the framework will be updated and adapted periodically to remain fit for purpose, building on the increasing number of evidence and use cases collected from across OECD countries.

The preliminary framework acknowledges the importance of emerging global standards and international collaboration to address the existing policy challenges. In particular, it is developed in the context of ongoing AI work at the OECD, such as the *Recommendation on AI* (OECD, 2023^[21]) and the *Framework of Tools for Trustworthy AI* (OECD, 2021^[4]).

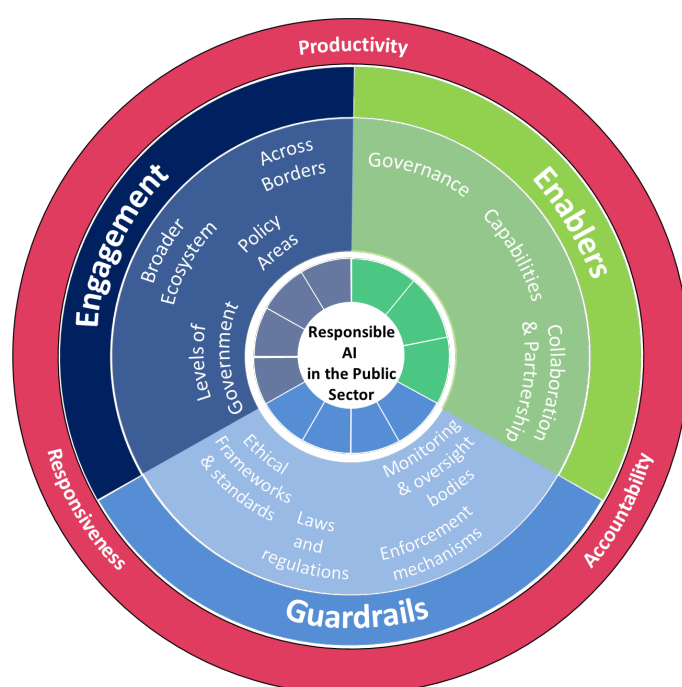
The framework is organised around three policy questions and four policy measures, as detailed in Table 2 and represented in Figure 1.

- The policy questions address three crucial aspects of public policy design:
 - *what* actions to take, *who* to engage, and *why* these actions are necessary. The application of the framework supports in *how* to implement actions, providing guidance for governments on the adoption of a holistic approach to a trustworthy AI in the public sector.
- The policy questions are complemented by policy options and actions:
 - on *what* actions - governments can deploy enablers in the fields of governance, capabilities, and collaborations and partnerships, and can develop a set of guardrails to guide, monitor and keep public institutions accountable.
 - on *who* to engage - governments can develop engagement strategies involving a wide range of stakeholders listed in the description column in Table 2.
 - on *why* - governments can consider three potential impacts outlined in Table 1. Understanding the use of AI in the public sector, in order to align AI usage with relevant public policy objectives.

Table 2. Policy questions and measures - framework for trustworthy use of AI in the public sector

Policy question	Policy measure	Description
What concrete policy actions and tools can governments develop to address existing constraints for a trustworthy use of AI in the public sector?	Enablers	Policy actions to establish a solid enabling environment and unlock the full-scale adoption of AI in the public sector. These include governance, capabilities (e.g. infrastructure, data, skills and talent), collaborations and partnerships in areas where policy makers currently identify constraints and shortcomings.
	Guardrails	Policy tools that governments can consider developing for a responsible, trustworthy and human-centred use of AI in the public sector. These may include soft laws and guidance as standards; policies; laws and regulations; enforcement mechanisms; oversight bodies (beyond the executive); monitoring bodies (within the executive); or collective agreements with public sector workforce.
Who should governments engage when developing and implementing the enablers and guardrails for the trustworthy use of AI in the public sector?	Engagement	Different stakeholders that need to be engaged in building the foundations for a responsible use of AI in the public sector. Various actors across the public sector (e.g. ministries, civil servants, sub-national governments), in the broader ecosystem and beyond national jurisdictions would need to be engaged through targeted actions to effectively address policy challenges related to the use of AI in the public sector.
What impact should the public sector strive to achieve when using trustworthy AI?	Impact	In connection with the AI uses cases taxonomy (see Table 1), AI in the public sector should strive for increasing productivity, responsiveness, and accountability.

Source: Authors.

Figure 1. Preliminary framework for trustworthy use of AI in the public sector

Source: Authors.

5 Conclusions

The use cases analysed in this paper show there is a clear potential for AI, if managed responsibly with the necessary safeguards, to positively transform how governments operate, formulate policies, deliver services and remain accountable. There is also growing awareness of the risks associated with the use of AI in the public sector. In recent years governments have undertaken a wide range of initiatives to create an enabling environment to responsibly govern with AI. The initiatives range from establishing new institutional arrangements and strategies to secure political and public support, to developing dedicated policies to effectively prevent and address failures or mitigate risks. Investments in capabilities and monitoring mechanisms are also acknowledged as critical tools for effectively deploying and overseeing the responsible use of AI. Nonetheless, there is a need for a more comprehensive, consistent and shared approach across public sectors.

There are alternative policy options governments can consider to establish an enabling environment for a trustworthy use of AI in the public sector. Building on the use cases analyses what seems to emerge is a commonality in the framing of the questions and the identification of the policy measures - enablers, guardrails, engagement – that governments can take to deliver the expected impacts. Section four of this paper proposes some actions that can be adopted at the national level, to guide public sectors in their pursuit of a safe, secure, and trustworthy use of AI. It provides also a common language for collecting evidence on successful and less positive experiences, fostering peer-learning and enhancing co-operation globally in addressing common challenges.

Looking ahead, understanding, promoting and enabling the *positive* aspects of using AI, rather than only preventing the *negative* ones, will remain a priority, since focusing mainly on risks might deter the deployment of high-benefit, low-risk uses of AI to improve public services. The systematic collection, documentation, and analysis of use cases on AI in the public sector will be needed to monitor trends on policy options across countries. More and better evidence of the impact of AI on governments will help ensure it is used for optimal impact. Easy access to such evidence, as well as information on policies, practices, and of AI systems used in the public sector – for example, via a global repository – could foster a structured dialogue and exchanges among countries.

The promotion of trustworthy use of AI will also require indicators and qualitative evidence on implementation and impacts. Finally, a trustworthy use of AI is the result of a collective effort that demands multistakeholder engagement and dialogue across policy sectors and beyond national borders to share lessons and jointly explore policy options as new challenges and opportunities emerge.

References

- Ada Lovelace Institute, AI Now Institute and Open Government Partnership (2021), *Algorithmic Accountability for the Public Sector*, <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>. [30]
- Berryhill, J. et al. (2019), “Hello, World: Artificial intelligence and its use in the public sector”, *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>. [24]
- Broecke, S. (2023), “Artificial intelligence and labour market matching”, *OECD Social, Employment and Migration Working Papers*, No. 284, OECD Publishing, Paris, <https://doi.org/10.1787/2b440821-en>. [12]
- Datta, P. et al. (2023), “GiusBERTo: Italy’s AI-Based Judicial Transformation: A Teaching Case”, *Communications of the Association for Information Systems*, Vol. 53/1, <https://aisel.aisnet.org/cais/vol53/iss1/33> (accessed on 13 October 2023). [11]
- Farrell, E. et al. (2023), *Artificial Intelligence for the Public Sector*, Publications Office of the European Union, <https://doi.org/doi:10.2760/91814>. [2]
- Future of Life Institute (2024), *AI Act Overview*, https://artificialintelligenceact.eu/wp-content/uploads/2024/01/AI-Act-Overview_24-01-2024.pdf (accessed on 10 April 2024). [18]
- González, F., T. Ortiz and R. Sánchez Ávalos (2020), *IA Responsable: Manual técnico: Ciclo de vida de la inteligencia artificial*, Inter-American Development Bank, <https://doi.org/10.18235/0002876>. [3]
- Lorenz, P., K. Perset and J. Berryhill (2023), “Initial policy considerations for generative artificial intelligence”, *OECD Artificial Intelligence Papers*, No. 1, OECD Publishing, Paris, <https://doi.org/10.1787/fae2d1e6-en>. [29]
- Manzoni, M. et al. (2022), *AI Watch Road to the adoption of Artificial Intelligence by the Public Sector: A Handbook for Policymakers, Public Administrations and Relevant Stakeholders*, Publications Office of the European Union, <https://publications.jrc.ec.europa.eu/repository/handle/JRC129100>. [20]
- OECD (2024), “2023 OECD Digital Government Index: Results and key findings”, *OECD Public Governance Policy Papers*, No. 44, OECD Publishing, Paris, <https://doi.org/10.1787/1a89ed5e-en>. [10]
- OECD (2024), *Anti-Corruption and Integrity Outlook 2024*, OECD Publishing, Paris, <https://doi.org/10.1787/968587cd-en>. [9]
- OECD (2023), “2023 OECD Open, Useful and Re-usable data (OURdata) Index: Results and key findings”, *OECD Public Governance Policy Papers*, No. 43, OECD Publishing, Paris, <https://doi.org/10.1787/a37f51c3-en>. [19]
- OECD (2023), *Global Trends in Government Innovation 2023*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/0655b570-en>. [25]

- OECD (2023), *Government at a Glance 2023*, OECD Publishing, Paris, [26]
<https://doi.org/10.1787/3d5c5d31-en>.
- OECD (2023), "Leveraging artificial intelligence for proactive delivery of public policies and services", in *Government at a Glance 2023*, OECD Publishing, Paris, [27]
<https://doi.org/10.1787/5702d250-en>.
- OECD (2023), *Recommendation of the Council on Artificial Intelligence (amended)*, [21]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.
- OECD (2023), "The state of implementation of the OECD AI Principles four years on", *OECD Artificial Intelligence Papers*, No. 3, OECD Publishing, Paris, [17]
<https://doi.org/10.1787/835641c9-en>.
- OECD (2022), *Digital Government Review of Luxembourg: Towards More Digital, Innovative and Inclusive Public Services*, OECD Digital Government Studies, OECD Publishing, Paris, [31]
<https://doi.org/10.1787/b623803d-en>.
- OECD (2022), "OECD Framework for the Classification of AI systems", *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, [7]
<https://doi.org/10.1787/cb6d9eca-en>.
- OECD (2022), *The Protection and Promotion of Civic Space: Strengthening Alignment with International Standards and Guidance*, OECD Publishing, Paris, [22]
<https://doi.org/10.1787/d234e975-en>.
- OECD (2022), "Transforming public governance for digital democracy", in *Building Trust and Reinforcing Democracy: Preparing the Ground for Government Action*, OECD Publishing, Paris, [13]
<https://doi.org/10.1787/01b73275-en>.
- OECD (2021), *Countering Public Grant Fraud in Spain: Machine Learning for Assessing Risks and Targeting Control Activities*, OECD Public Governance Reviews, OECD Publishing, Paris, [14]
<https://doi.org/10.1787/0ea22484-en>.
- OECD (2021), *Data-Driven, Information-Enabled Regulatory Delivery*, OECD Publishing, Paris, [38]
<https://doi.org/10.1787/8f99ec8c-en>.
- OECD (2021), *G20 Compendium on the Use of Digital Tools for Public Service Continuity: Report for the G20 Digital Economy Task Force, Trieste, Italy, August 2021*, OECD Publishing, Paris, [23]
<https://doi.org/10.1787/6f800fd5-en>.
- OECD (2021), *G20 survey on Agile approaches to the regulatory governance of innovation: Report for the G20 Digital Economy Task Force, Trieste, Italy, August 2021*, OECD Publishing, Paris, [36]
<https://doi.org/10.1787/f161916d-en>.
- OECD (2021), *Good Practice Principles for Data Ethics in the Public Sector*, [34]
<https://www.oecd.org/gov/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.pdf>.
- OECD (2021), *Recommendation of the Council for Agile Regulatory Governance to Harness Innovation*, [37]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0464>.
- OECD (2021), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, [33]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463>.
- OECD (2021), "State of implementation of the OECD AI Principles: Insights from national AI [5]

- policies”, *OECD Digital Economy Papers*, No. 311, OECD Publishing, Paris, <https://doi.org/10.1787/1cd40c44-en>.
- OECD (2021), “The OECD Framework for digital talent and skills in the public sector”, *OECD Working Papers on Public Governance*, No. 45, OECD Publishing, Paris, <https://doi.org/10.1787/4e7c3f58-en>. [28]
- OECD (2021), “Tools for trustworthy AI: A framework to compare implementation tools for trustworthy AI systems”, *OECD Digital Economy Papers*, No. 312, OECD Publishing, Paris, <https://doi.org/10.1787/008232ec-en>. [4]
- OECD (2020), *Building Capacity for Evidence-Informed Policy-Making: Lessons from Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/86331250-en>. [8]
- OECD (2019), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, <https://doi.org/10.1787/059814a7-en>. [32]
- OECD/CAF (2022), *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/1f334543-en>. [6]
- Oliveira, T., A. Monteiro da Rocha and M. Scatolino de Rezende (2022), “Alice: Desafios, resultados e perspectivas da ferramenta de auditoria contínua de compras públicas governamentais com uso de inteligência artificial”, *Revista da CGU*, Vol. 14/26, <https://doi.org/10.36428/revistadacgu.v14i26.530>. [15]
- Tangi, L. et al. (2022), *AI Watch. European landscape on the use of Artificial Intelligence by the Public Sector*, Publications Office of the European Union, <https://doi.org/doi:10.2760/39336>. [35]
- Ubaldi, B. et al. (2019), “State of the art in the use of emerging technologies in the public sector”, *OECD Working Papers on Public Governance*, No. 31, OECD Publishing, Paris, <https://doi.org/10.1787/932780bc-en>. [1]
- Wong, J. et al. (2022), “Key challenges for the participatory governance of AI in public administration”, *SocArXiv*, <https://doi.org/10.31219/OSF.IO/PDCRM>. [16]

End notes

¹ Throughout this paper, artificial intelligence (AI) will refer to a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment (OECD, 2023^[21]).

² The OECD has previously documented the use of AI in the public sector and its policy implications. This paper builds upon policy analysis on AI in the public sector developed throughout conceptual pieces

(Ubaldi et al., 2019^[1]; Berryhill et al., 2019^[24]), a compendium of best practices for the G20 (OECD, 2021^[23]), reviews and standards on data-informed approaches to rule making and enforcement, including the use of AI solutions (OECD, 2021^[38]; OECD, 2021^[36]; OECD, 2021^[37]), national and regional reviews (OECD, 2022^[31]; OECD/CAF, 2022^[6]), global government innovation trends reporting (OECD, 2023^[25]), and more recently as part of ‘Government at a Glance 2023’ (OECD, 2023^[27]) and the 2023 editions of the *OECD Digital Government Index* (OECD, 2024^[10]) and *OURdata Index* (OECD, 2023^[19]).

³ <https://ai-magazine.vngrealisatie.nl/nijmegen/>

⁴ <https://www.boost.ai/case-studies/how-conversational-ai-is-helping-norways-citizens-with-covid>

⁵ For further information, see: <https://tc.canada.ca/en/programs/pre-load-air-cargo-targeting-pact> and <https://canadagazette.gc.ca/rp-pr/p1/2023/2023-03-18/html/reg5-eng.html>

⁶ <https://www.eff.org/deeplinks/2022/05/movement-ban-government-use-face-recognition>

⁷ <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>

⁸ https://www.japan.go.jp/kizuna/2024/02/hiroshima_ai_process.html

⁹ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

¹⁰ <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>

¹¹ <https://news.un.org/en/story/2024/03/1147831>

¹² <https://innovazione.gov.it/notizie/articoli/en/g7-ministerial-declaration-on-industry-technology-and-digital/>

¹³ For a comprehensive taxonomy of AI systems, refer to the *OECD Framework for the Classification of AI Systems* (OECD, 2022^[7]).

¹⁴ A non-exhaustive list of AI use cases in the public sector across OECD countries is available upon request. Interested parties may contact eleaders@oecd.org for access to the data. Also refer to the European Commission’s Public Sector Tech Watch, which includes a systematic collection of more than a thousand public sector use cases across the European Union: <https://joinup.ec.europa.eu/collection/public-sector-tech-watch/cases>.

¹⁵ “Generative artificial intelligence (AI) systems create new content in response to prompts based on their training data” (Lorenz, Perset and Berryhill, 2023, p. 8^[29]).

¹⁶ <https://www.france-services.gouv.fr/actualites/experimentation-dun-modele-dassistance-france-services-IA>

¹⁷ https://www.bolagsverket.se/omoss/press/pressmeddelanden/pressmeddelanden2023/aimodellsortera_rmyndighetensinkorg.4936.html

¹⁸ <https://www.gov.uk/government/news/government-to-harness-the-power-of-ai-to-improve-public-project-delivery-under-new-framework>

¹⁹ <https://oecd-opsi.org/innovations/wastelands-a/>

²⁰ https://www.paris-saclay.com/fileadmin/documents/1._L_agglo/Espace_Presse/Paris-Saclay_Twin_2030.pdf

²¹ <https://oecd-opsi.org/innovations/auroraai/>

²² “Personalisation: Develops a profile of an individual and then learns and adapts to that individual over time. The output is usually a ranking, e.g. a search engine ranking.” (OECD, 2022, p. 50^[7]).

²³ <https://neyimvar.gov.tr/giris>

²⁴ <https://www.kratid.ee/kasutuslood-kratid> and <https://joinup.ec.europa.eu/collection/public-sector-tech-watch/cases>

²⁵ See for instance the Catalogue of Tools & Metrics for Trustworthy AI of the OECD AI Policy Observatory (<https://oecd.ai/en/catalogue/overview>).

²⁶ “Algorithmic accountability is defined as ensuring that those that build, procure and use algorithms are eventually answerable for their impacts” (Ada Lovelace Institute, AI Now Institute and Open Government Partnership, 2021, p. 4^[30]). Its importance has been globally recognised in policy fora, for instance through the *OECD Recommendation of the Council on Artificial Intelligence*, which establishes that “AI actors should be accountable for the proper functioning of AI systems”. Algorithmic accountability relies on concepts of algorithmic transparency, which refers to the availability and accessibility of information on these algorithms, such as the data used, the source code, the algorithm’s goals and usage, and compliance with certain standards.

²⁷ <https://www.regjeringen.no/en/dep/dfd/id810/>

²⁸ <https://digital-strategy.ec.europa.eu/en/policies/ai-office>

²⁹ https://algorithmic-transparency.ec.europa.eu/index_en

³⁰ <https://espanadigital.gob.es/actualidad/aprobado-el-estatuto-de-la-agencia-espanola-de-supervision-de-la-inteligencia-artificial>

³¹ <https://democracy-technologies.org/participation/belgium-citizens-assembly-ai/>

³² <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf>

³³ <https://www.dta.gov.au/blogs/ai-government-taskforce-examining-use-and-governance-ai-aps>

³⁴ <https://www.conseil-etat.fr/publications-colloques/etudes/intelligence-artificielle-et-action-publique-construire-la-confiance-servir-la-performance>

³⁵ <https://www.fonction-publique.gouv.fr/toutes-les-publications/strategie-dusage-de-lintelligence-artificielle-en-matiere-de-gestion-des-ressources-humaines-dans-la-fonction-publique-de-letat>

³⁶ <https://oecd.ai/en/wonk/documents/uruguay-ai-strategy-for-the-digital-government-2019>

- ³⁷ <https://inteligenciaartificial.gov.co/marco-eticos/>
- ³⁸ <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html>
- ³⁹ https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240419_9.pdf
- ⁴⁰ https://guides.etalab.gouv.fr/algorithmes/guide/#_3-le-cadre-juridique-applicable
- ⁴¹ <https://www.gov.ie/en/publication/2127d-interim-guidelines-for-use-of-ai/>
- ⁴² https://www.berenschot.nl/media/vf5fj1mw/handreiking_governance-voor-een-verantwoorde-toepassing-van-algoritmen.pdf
- ⁴³ <https://oecd-opsi.org/innovations/algorithmic-transparency-standard/>
- ⁴⁴ <https://citiesfordigitalrights.org/>
- ⁴⁵ <https://aiforgood.itu.int/helsinki-and-amsterdam-launch-ai-registers-to-detail-city-systems/>
- ⁴⁶ <https://gouai.cidob.org/>
- ⁴⁷ <https://citiesfordigitalrights.org/digitalrightsgovernance>
- ⁴⁸ <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>
- ⁴⁹ <https://women-gender-equality.canada.ca/en/gender-based-analysis-plus.html>
- ⁵⁰ <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf>
- ⁵¹ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- ⁵² <https://www.ria.ee/en/state-information-system/people-centred-data-exchange/consent-service>
- ⁵³ <https://e-estonia.com/data-tracker-build-citizen-trust/>
- ⁵⁴ See *Bürokratt* virtual assistant, <https://www.ria.ee/en/state-information-system/machine-learning-and-language-technology-solutions/burokratt>
- ⁵⁵ <https://www.elementsofai.com/eu2019fi>
- ⁵⁶ <https://oecd-opsi.org/innovations/ai-certification-ireland/>
- ⁵⁷ https://ai-watch.ec.europa.eu/countries/estonia/estonia-ai-strategy-report_en
- ⁵⁸ See IEEE P3119 Standard for the Procurement of Artificial Intelligence and Automated Decision Systems from 2021 (<https://standards.ieee.org/ieee/3119/10729/>).

⁵⁹ <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai>

⁶⁰ https://assets.publishing.service.gov.uk/media/60b356228fa8f5489723d170/Guidelines_for_AI_procurement.pdf

⁶¹ <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/list-interested-artificial-intelligence-ai-suppliers.html>

⁶² <https://www.turing.ac.uk/research/research-programmes/public-policy>

⁶³ <https://playbook.cio.gov/techfar/>

⁶⁴ [AIM: The OECD AI Incidents Monitor, an evidence base for trustworthy AI - OECD.AI](#)

⁶⁵ <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>

⁶⁶ <https://algoritmes.overheid.nl/>

⁶⁷ See for instance “Auditing machine learning algorithms”, a white paper for public auditors released by the Supreme Audit Institutions of Finland, Germany, the Netherlands, Norway and the UK: <https://www.auditingalgorithms.net/>. On work being developed by data protection agencies, see for instance the Dutch Data Protection Authority’s “Periodic insight into the risks and effects of the use of AI & algorithms in the Netherlands” (<https://www.autoriteitpersoonsgegevens.nl/uploads/2024-01/AI%20%26%20Algorithmic%20Risks%20Report%20Netherlands%20-%20winter%202023%202024.pdf>). On work being developed by ombudspersons, see for instance The European Ombudsman’s inquiry to the European Commission on how it decides on and uses AI in its decision making (<https://www.ombudsman.europa.eu/en/news-document/en/183633>).

⁶⁸ <https://www.riksrevisjonen.no/contentassets/9e1061c5f7fe4678b4b93e5d1e5bb611/annual-report-2022.pdf>

⁶⁹ <https://www.riksrevisjonen.no/globalassets/om-riksrevisjonen/auditingtobenefitthesocietyoftomorrow.pdf>

⁷⁰ <https://www.nao.org.uk/reports/use-of-artificial-intelligence-in-government/>

⁷¹ <https://espanadigital.gob.es/lineas-de-actuacion/agencia-nacional-de-supervision-de-la-inteligencia-artificial> and https://boe.es/diario_boe/txt.php?id=BOE-A-2023-18911

⁷² <https://www.gao.gov/products/gao-21-519sp>

⁷³ <https://www.rekenkamer.nl/onderwerpen/algoritmes/algoritmes-toetsingskader/>

⁷⁴ <https://oecd-opsi.org/blog/chile-algorithmic-transparency/>

⁷⁵ <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf>

