OECD

Governing with Artificial Intelligence

The State of Play and Way Forward in Core Government Functions

Governing with Artificial Intelligence

THE STATE OF PLAY AND WAY FORWARD IN CORE GOVERNMENT FUNCTIONS

OECD

BETTER POLICIES FOR BETTER LIVES

This work was approved and declassified by the Public Governance Committee on 05/09/2025.

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.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Note by the Republic of Türkiye
The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Türkiye recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Türkiye shall preserve its position concerning the "Cyprus issue".

Note by all the European Union Member States of the OECD and the European Union
The Republic of Cyprus is recognised by all members of the United Nations with the exception of Türkiye. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

Photo credits: Cover © cofotoisme/Getty Images.

Foreword

In recent years, governments worldwide have made significant strides in the digitalisation of the public sector, accelerated by the challenges posed by the COVID-19 pandemic. Despite these advancements, many obstacles remain. Today, the focus is on harnessing digital technologies to better meet the needs of citizens — creating efficiencies to deliver more value for taxpayers' money; providing tailored, accessible and inclusive public services; and improving communication and engagement with citizens.

The importance of these efforts cannot be overstated. Enhanced government efficiency and effectiveness are crucial, but the trust citizens place in their governments is equally important. Trust in government is a key factor in the success of digital initiatives. Conversely, while OECD data from 2023 shows that only 39% of people have moderately high or greater trust in national government, reliable, responsive and fair public services can enhance trust in government.

Artificial intelligence (AI) is becoming a significant component of the digital government journey, offering substantial benefits in various areas such as automation, anomaly detection and improved decision making. For example, AI-powered chatbots are being used to answer citizen queries and assist with form filling. In disaster management, AI is helping to anticipate natural disasters and speed up response efforts. In tax administration, AI is being used for fraud detection. Overall, the use of AI in government can improve government productivity, responsiveness and accountability.

However, AI adoption in government trails behind that of the private sector. Governments face unique contexts and challenges that hinder the rapid uptake of AI, including skills shortages, outdated legacy systems, data availability and a financially constrained environment, as well as higher requirements regarding privacy, transparency and representation of various groups.

In this fast-paced digital landscape, learning from experience is the best way to make progress. While AI maturity is not yet prevalent in governments, there are many examples of AI applications from which lessons can be drawn. The OECD is committed to supporting governments on their AI journey by facilitating the sharing of experiences and insights. This report, built on the analysis of dozens of governance approaches and 200 AI use cases, extensive research and discussions with governments serves as a foundation for an ongoing workstream on "Governing with AI" (https://oecd.ai/gov). It aims to provide governments with the necessary elements for effective AI use and to identify areas where AI is having an impact and where gaps remain. Future efforts will build on the growing evidence base of policies and use cases, seeking to further assist governments putting in place the enablers, guardrails and engagement mechanisms needed for a strategic and trustworthy approach to AI.

The OECD acknowledges the efforts of many countries in accelerating the adoption of AI in government beyond the data and insights analysed for this report. Their examples have been instrumental in shaping our understanding of AI's role in transforming government.

This report is part of the OECD Horizontal Project on Thriving with AI: Empowering Economies and Societies.

Acknowledgements

- **Justice administration and access to justice**: Tatyana Teplova, Maaike de Langen and Mariane Piccinini Barbieri of GOV's Global Partnerships, Inclusion and Justice Division (GPIJ) and Paola Gálvez Callirgos, consultant. Contributions were also made by Celine Cojocar (GOV/GPIJ), Aitor Cubo, Fabiola Solino Díaz and Alejandro Fernández Muñoz (Spain); and Tomislav Boršić (Croatia).

Table of contents

FIGURES

TABLES

Executive summary

Artificial intelligence (AI) is one of the most transformative forces of the 21st century, and it is becoming an integral part of digital government worldwide. Governments' use of AI can facilitate automated and tailored internal processes and public services, foster better decision making and forecasting, improve fraud detection and improve public servants' job quality and learning – all with tangible impacts. For example, The Alan Turing Institute estimates that the AI could automate 84% of repetitive public service transactions in the United Kingdom, saving the equivalent of 1 200 person-years of work annually. Despite its promise, government AI use trails the private sector.

Key findings: How AI can serve citizens

The OECD has conducted in-depth research of AI in 11 core functions of government across 200 use cases. The results suggest that AI is most prevalent in terms of total use cases in public service and justice functions and civic participation, with relatively less use seen in policy evaluation, tax administration and civil service reform. In between are public procurement, financial management, fighting corruption and promoting public integrity, and regulatory design and delivery. Possible explanations for this distribution include that some functions encompass a wider variety of uses (public services) while others are more narrow (civil service reform, tax administration). Also, some face more regulatory constraints (e.g. tax administration, given rules on using tax data), while some face fewer implementation challenges and can mature faster (civic participation). In some functions, such as justice administration, public demands and growing transactions backlogs may precipitate AI adoption as an opportunity to tackle urgent challenges.

AI's use is more prevalent in internal operations and public service delivery, but less prominent in government oversight. Less use is also seen in policymaking, consistent with previous OECD analysis. Use cases often rely on classic rules-based approaches or established machine learning (ML) techniques, with generative AI (GenAI), including large language models (LLMs), being less common. In terms of benefits, the largest share of cases seeks to promote automated, streamlined and tailored processes and services; followed by better decision making and forecasting; and enhanced accountability and anomaly detection. A few cases seek to unlock new opportunities for external stakeholders (e.g. citizens, businesses) through access to government-provided AI systems, but further efforts may be warranted.

Risks for AI use in government

There is no such thing as risk-free AI adoption. Unlocking AI's benefits requires mitigating its risks. Biased algorithms can result in adverse outcomes; AI misuse can infringe on or prevent the free exercise of human rights; insufficient transparency, explainability and public understanding of AI can erode accountability and cause public resistance; and the over-reliance on AI can widen digital divides and allow systemic errors to propagate, weakening citizen trust in government. Such risks may be amplified in countries lacking mechanisms to guarantee the exercise, protection and promotion of human rights, or result from AI misuse

by individual public servants. Public service workforce displacement could also occur if governments seek to replace rather than augment civil servants' capabilities.

Governments' failure to leverage AI also represents risk, resulting in missed opportunities to yield benefits and widening the gap between public and private sector capacities. They will need to adopt AI if they want to meet increasing citizen demands and strengthen trust in government. Ignoring AI transformation or waiting for all unknowns to be resolved relegates government to being a technology-taker rather and an option-shaper, incurring significant costs and disadvantages. If governments do not bolster internal AI capacities soon, they may struggle to ever catch up.

Governments also face AI implementation challenges

Challenges in scaling up successful AI applications means government AI initiatives often remain in pilot phases. Skills gaps and difficulties in obtaining and sharing quality data are encountered across government functions. Moreover, although strategies for AI in government are common, a lack of concrete guidance hinders their transformation into practice. These factors compound risk aversion, hindering governments' ability to innovate with AI. Furthermore, insufficient monitoring and evaluation mechanisms restrict their ability to gauge progress, detect risks and demonstrate return on investment. Financial costs are also a common challenge.

Some challenges are more prevalent in some functions than others. For instance, tax administration faces complex laws and rules around tax processes and data, whereas public procurement struggles with a lack of established rules around AI. Finally, the use of AI in functions such as public financial management is constrained by outdated legacy technology infrastructure unsuitable for AI development or use.

How governments can ensure their use of AI is trustworthy

To reap the benefits of AI in government while mitigating its risks and overcoming implementation challenges, governments need to put in place:

- **Enablers** to facilitate trustworthy adoption, including governance, data, digital infrastructure, skills, financial investments, agile procurement processes and capacities to partner with non-governmental actors.
- **Guardrails** to guide the use of AI, including rules and policies, guidance and frameworks, transparency and accountability mechanisms that span the AI system lifecycle, and oversight and advisory bodies to guide and evaluate efforts.
- **Engagement** approaches to shape user-centred and responsive approaches, including mechanisms to engage with key stakeholders, including the public, civil society and businesses.

More action is needed to invest in and adopt trustworthy AI in government, but existing approaches provide lessons and inspiration

To the extent possible, the OECD encourages governments to prioritise high-benefit, low-risk applications of AI, especially when building an initial level of maturity. Most, however, lack the processes for holistic measurement of potential or realised results — spending efficiency, service quality, potential harms — that would allow them to make these determinations. Addressing this should be a priority for governments, ensuring AI implementations are transparent, fair and secure.

Many government AI efforts are in their infancy, but some are yielding valuable lessons. The OECD is committed to expanding an evidence base of what works through data collection and analysis, with a focus on how governments can leverage trustworthy AI to deliver public value.

Acronyms and abbreviations

AGI	Artificial general intelligence
AI	Artificial intelligence
AIA	Algorithmic impact assessment
AIG	Anticipatory Innovation Governance
AIM	OECD AI Incidents Monitor
API	Application programming interface
APS	Australian Public Service
ATI	Access to information
AU	African Union
BLOOM	BigScience Large Open-science Open-access Multilingual Language Model
CIAO	Chief Artificial Intelligence Officer
CoE	Council of Europe
COMPAS	Correctional Offender Management Profiling for Alternative Sanctions (United States)
CSO	Civil society organisation
DGI	OECD Digital Government Index
DPI	Digital public infrastructure
DSIT	Department for Science, Innovation and Technology (United Kingdom)
E-Leaders	OECD Working Party on Senior Digital Government Officials
EC	European Commission
EU	European Union
FMIS	Financial management information system
FTE	Full-time equivalent
GDP	Gross Domestic Product
GDPR	General Data Protection Regulation (European Union)
GenAI	Generative artificial intelligence
GPAI	Global Partnership on Artificial Intelligence (GPAI)
HR	Human resource
HRM	Human resource management
i.AI	Incubator for AI (United Kingdom)
IAPR	Independent Authority for Public Revenue (Greece)
ITTI	Inventory of Tax Technology Initiatives (OECD)
JRC	Joint Research Centre (European Commission)

LLM	Large language model
ML	Machine learning
NER	OECD Network of Economic Regulators
NIST	National Institute of Standards and Technology (United States)
NLP	Natural Language Processing
OCR	Optical character recognition
OECD	Organisation for Economic Co-operation and Development
OGD	Open government data
OMB	Office of Management and Budget (United States)
PAC	Public Accounts Committee (Parliament of the United Kingdom)
Palkeet	Government Shared Services Centre for Finance and Human Resources (Finland)
PET	Privacy-enhancing technology
PFM	Public financial management
PPP	Public-private partnership
R&D	Research and development
RaC	Rules as code
RAG	Retrieval-augmented generation
RCT	Randomised control trial
RIA	Regulatory impact assessment
ROI	Return on investment
RPA	Robotic process automation
SAI	Supreme Audit Institution
TCU	Federal Court of Accounts (Brazil)
SDG	Sustainable Development Goal
UN	United Nations

Note: Terms are generally only included if the abbreviation or acronym is mentioned outside of context of where it is originally defined. Abbreviations for country names are not included.

1 How artificial intelligence is accelerating the digital government journey

This chapter explains how artificial intelligence (AI) can accelerate the digital government journey. It situates government as a developer and user of AI, going beyond traditional investor and regulator roles. The chapter groups opportunities – productivity (efficiency and effectiveness), responsiveness and accountability – across the policy cycle, and stresses prerequisites in data and information management. It also outlines government-specific risks and risks of inaction, within emerging regulatory approaches, and closes with a vision for trustworthy AI in government.

Key messages

- Artificial intelligence (AI) has the potential to reshape industries, economies, governments and societies. Yet, its progress in the government has been limited.
- AI can help governments in three key opportunity areas: productivity, responsiveness and accountability.
- At each stage of the policy cycle, AI can bring highly complementary benefits:
  o automating mundane and repetitive tasks
  o improving productivity in analytical or creative tasks
  o tailoring services to address personalised citizen needs
  o tailoring approaches to strengthen the civil service
  o enhancing decision-making and sense-making of the present
  o better forecasting of the future
  o improving information management and accessibility
  o detecting improper transactions and assessing integrity risks
  o enabling non-governmental actors to understand and engage with government and promote accountability
  o unlocking opportunities for external stakeholders through AI as a good for all.
- These benefits are not mutually exclusive and can be categorised into four broad areas:
  o automated, streamlined and tailored processes and services
  o better decision-making, sense-making and forecasting
  o enhanced accountability and anomaly detection
  o unlocking opportunities for external stakeholders.
- Governments should manage the risks of AI that are specific to government use, which are: ethical risks, operational risks, exclusion risks, public resistance risks and risks of inaction.
- A vision of a future where governments successfully develop and adopt trustworthy AI for systematic transformation of government processes and services is beginning to emerge.

The digital government journey

Digital government is essential to transforming processes and services in ways that improve the public sector's responsiveness and reliability and bring governments closer to their people. Since the adoption of the OECD Recommendation on Digital Government Strategies (2014[1]), the OECD has been promoting digital government in OECD member countries and beyond, supporting them in their efforts to achieve government digital maturity. Digitally mature governments recognise that technology is a strategic driver not only to improve efficiency, but also to make policies more effective and governments more open, accountable, innovative, participatory and trustworthy.

The COVID-19 pandemic underscored the importance of digital technologies and data in building economic and social resilience through strategic, agile and innovative government approaches. While the pandemic and the multidimensional crisis it provoked disrupted governments, it also offered an opportunity to revisit strategic approaches on the use of digital tools and data to improve the delivery of public services. Faced with no alternative, governments compressed years' worth of technological advancements into weeks and months. Deploying technology solutions at scale enabled governments to continue operating in times of crisis, and secured the timely provision of services to citizens and businesses (OECD, 2020[2]; [3]). Where digital technologies or data were not used strategically or effectively, the crisis highlighted gaps and exacerbated challenges, which governments are working to address to this day.

Today, governments worldwide are facing decreasing levels of public trust (OECD, 2024[4]), while simultaneously experiencing growing and rapidly accelerating changes brought about by the digital age. In this time of fast-paced disruption — rapid technological evolution, changing societal needs, unexpected crises — it is crucial governments be capable and equipped to use digital technologies and data to increase productivity and resilience in their public administrations, and enhance the quality of public services.

Institutionalising digital government, with varying maturity levels

To unlock the potential of digital government, establishing the right institutional arrangements, coordination mechanisms and policy instruments is critical to sustaining the needed long-term transformations and overcoming changing political priorities. The OECD (2020[3]) Digital Government Policy Framework establishes six dimensions critical for establishing a digital government:

1. **Digital by design**: designing policies to enable the public sector to use digital tools and data in a coherent way when formulating policies or transforming public services.
2. **Data-driven**: developing the governance and enablers needed for data access, sharing and re-use across the public sector.
3. **Government as a platform**: deploying common building blocks such as guidelines, tools, data, digital identity and software to advance a coherent transformation of government processes and services across the public sector.
4. **Open by default**: openness beyond the release of open data, including efforts to encourage the use of technologies and data to communicate and engage with different actors.
5. **User-driven**: placing user needs at the core of the design and delivery of public policies and services, including through engagement with users and measurement of metrics to assess impact and satisfaction.
6. **Proactiveness**: anticipating the needs of users and service providers to deliver government services proactively.

The OECD Digital Government Index (DGI) benchmarks governments' maturity across these dimensions (Figure 1.1). In this figure, it is clear that some countries are further progressed in their journey towards

digital government maturity than others, and the broad array of OECD analysis on digital government illuminates that each country faces its own challenges in achieving maturity.[1]

Figure 1.1. OECD 2023 Digital Government Index, composite results by country
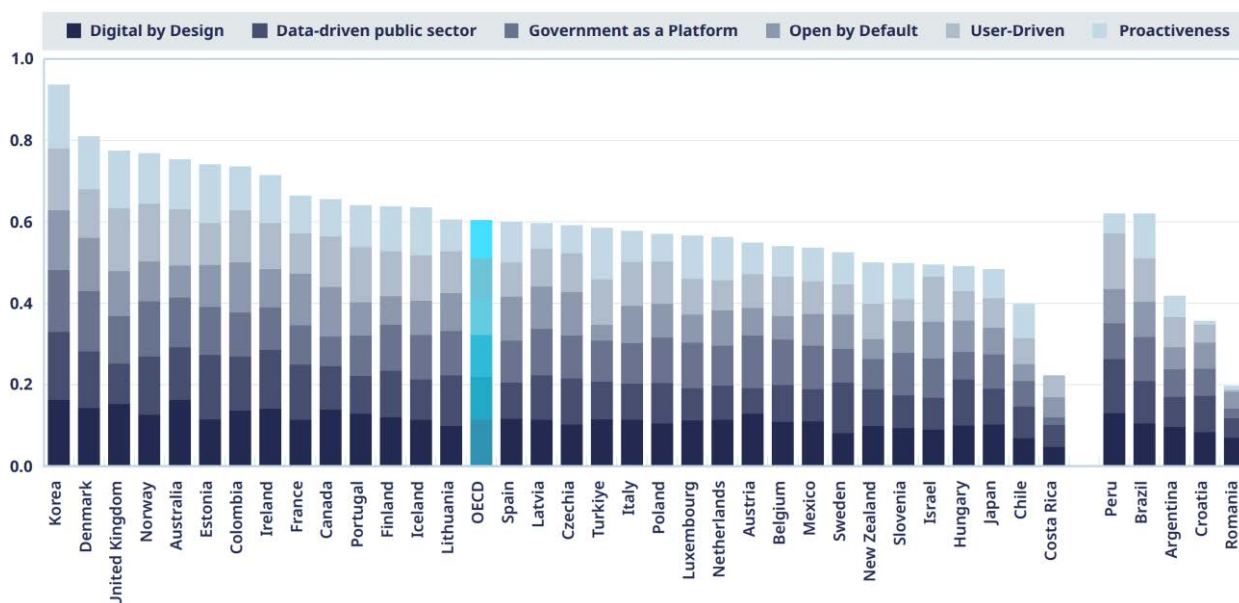
Note: Data for Germany, Greece, Slovakia, Switzerland and the United States (US) are not included.
Source: (OECD, 2024[5]).

AI's growing role in digital government

The OECD defines an AI system as:

> *"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."* – see OECD Explanatory Memorandum for further clarification *(2024[6]; [7])*.

Global AI discussions mainly focus on governments as AI regulators or investors, but significant opportunities exist for government as an AI developer and user. Not only do governments set national priorities, investments and regulations for AI, but they increasingly use it to design and implement policies and services. Although hype around AI has increased in recent years, governments are not new to using AI; there are thousands of government AI projects underway around the world.[2]

Since 2019, the OECD has been working to better understand AI's uses and implications in the unique context of government. This includes developing foundational pieces on the technical underpinnings of AI and its use and implications by and for government (2019[8]; [9]); targeted analysis on specific countries (2022[10]; 2024[11]; [12]); surfacing government innovation trends, which often involve AI;[3] and establishing a preliminary framework for AI in government (2024[13]). The OECD has also collected details on hundreds of initiatives for AI in government.[4]

The 2023 DGI highlights while some governments have been deploying a wide range of initiatives to enhance their capacity to use AI, implementation is still a challenge for most. At the time of the DGI's publishing, 70% of countries had used AI to improve internal governmental processes, while only 33% had used AI to enhance policy design and implementation. Although use is increasing, AI use in government has not yet made a transformative impact. The forthcoming 2025 DGI will include updated figures and more in-depth comparative analysis. It will also incorporate complementary qualitative evidence to further

inform how governments can implement the right enablers, safeguards, risk mitigation and engagement mechanisms to adopt trustworthy AI while monitoring for adverse effects.

Understanding AI's transformative potential

AI is one of the most transformative forces of the 21st century. It is reshaping industries, economies, governments and societies at an unprecedented pace. If governments and other AI actors are successful in seizing AI's benefits while mitigating its risks, AI experts and researchers envision a future in which AI contributes to scientific and medical breakthroughs, such as discovering new cancer treatments; catalyses productivity growth from a 1-7% rise in global gross domestic product (GDP) by 2033 to a 10-fold increase in the decades to come; eliminates poverty and reduces inequality; and helps address weather-related impacts and natural disasters (OECD, 2024[14]).

While AI has gained intense worldwide attention in recent years, AI research and development has been going on for over 70 years. Before taking a deeper look at AI use in government, it is useful to understand some of the background on AI and why it has recently become a topic of household discussion, as discussed in Box 1.1.

Box 1.1. The evolution of AI

The AI landscape has evolved significantly since the 1950s when British mathematician Alan Turing first posed the question of whether machines can think. For decades, **"rules-based" or "symbolic" AI systems** dominated research, using a series of "If-then" (If a condition, then an action) statements that, when taken together, would give the appearance of intelligent action. Such systems are limited and require significant human knowledge to programme the rules. They are still in use today, such as in robotic process automation (RPA) software bots that automate human-programmed tasks. Due to their limitations, some argue that rules-based systems and RPA should not be considered AI at all.

The 21st century saw breakthroughs in the branch of AI called **machine learning (ML)** that improved the ability of machines to make predictions from historical data. ML focuses on the development of systems that can learn and adapt without following explicit instructions imitating the way humans learn, gradually improving its accuracy, by using algorithms and statistical models to analyse and draw inferences from patterns in data. The "learning" process using machine-learning techniques is known as "training".

The application of ML techniques, the availability of large datasets, and faster and more powerful computing hardware have converged, dramatically increasing the capabilities, impact and availability of AI models and systems. Inspired by the human brain, **neural networks** are made up of layers of "neurons", known as "nodes", that process inputs with weights and biases to give specific outputs. A subset of algorithms in the area of neural networks — called deep neural networks (in the field of study and set of techniques called **deep learning**) — allows machine-based systems to "learn" from examples to make predictions or "inferences" based on the large amount of data processed during their training phase. Because of their complexity, it can be difficult to understand how they work or produce a given output.

Recent conceptual breakthroughs

In 2017, Google researchers introduced a type of neural network architecture called "**transformers**", which learn to detect how data elements — such as the words in this sentence — influence and depend on each other. Unlike previous neural networks, transformers can process inputs from a sequence, such as words of text, in parallel. This unleashed major progress by enabling AI developers to design

---

larger-scale language models with more parameters and greater efficiency. This contributed greatly to advances in **generative AI (GenAI), including large language models (LLMs)**, that can generate novel content and enable consumer-facing applications like advanced chatbots at people's fingertips.

For many, AI became "real" in 2022, the year that OpenAI's ChatGPT (Chat Generative Pre-Trained Transformer) became the fastest-growing consumer application in history. Transformers also contributed to the advent of **foundation models**, which are trained on large amounts of data that can be adapted (i.e. fine-tuned) and built upon to conduct a wide range of downstream tasks. Although transformers are often discussed, other approaches exist, especially for non-text (e.g. images, video, audio) generation, such as generative adversarial network (GANs) and diffusion models.

Most AI today is "narrow", but some argue more "general" forms of AI are emerging

Most AI today can be considered "**narrow**" (designed to perform a specific task), but some experts argue that foundation models are an early form of more "general" AI. This includes progress towards the hypothetical advent of **artificial general intelligence (AGI)** — a controversial concept that can be described as machines with human-level or greater intelligence across a broad spectrum of contexts. There is substantial debate and uncertainty among experts about when or if AGI might be developed, and what potential opportunities and challenges it may bring.

While some experts believe AGI will be developed at some point, emerging early forms of **"agentic" AI systems** — which can operate in a somewhat autonomous manner without the constant need for human guidance — hint at the potential for future systems that can handle more general tasks with minimal human input. For instance, LLM-based "agents" have already been developed to autonomously search the internet and interpret what they find on behalf of the user. Such systems are very early, comprising many limitations and risks, but further advancements may yield opportunities across sectors.

As AI systems become increasingly capable, many argue that humans should not defer decisions to machines but rather work in tandem, or in **machine-human collaboration,** using AI to assist decision-making.

Note: The OECD primer Hello World: Artificial intelligence and its use in the public sector (Berryhill et al., 2019[8]) provides details on the technical underpinnings and potential implications of AI.
Source: (OECD, 2024[15]), (OECD, 2024[14]), (OECD.AI, 2023[16]), (Lorenz, Perset and Berryhill, 2023[17]), (Berryhill et al., 2019[8]) (Cognitus, 2024[18]), (Purdy, 2024[19]), (NIST, 2025[20]), (Horvitz, 2014[21]), (Brizuela et al., 2025[22]), https://playbooks.aip.gov.sg/agentic-ai-primer.

Despite the hype, AI progress is limited

Data from the OECD.AI Policy Observatory demonstrates the boom in interest in AI in recent years. For example, Figure 1.2 shows significant growth in venture capital investments in AI over time. While interest is high, advisory firm Gartner's latest "hype cycle" puts GenAI just beyond its "peak of inflated expectations", and starting a descent into the "trough of disillusionment", "as business focus continues to shift from excitement around foundation models to use cases that drive ROI" (Gartner, 2024[23]). Still, it expects GenAI and some other forms of AI, such as AI supercomputing and the use of AI to support and enforce AI governance policies, trust, risk and security, to reach fuller productivity in just two to five years.

Figure 1.2. Venture capital investments in AI have grown over the years

Venture global capital investments in AI in USD millions by country from 2012 onwards
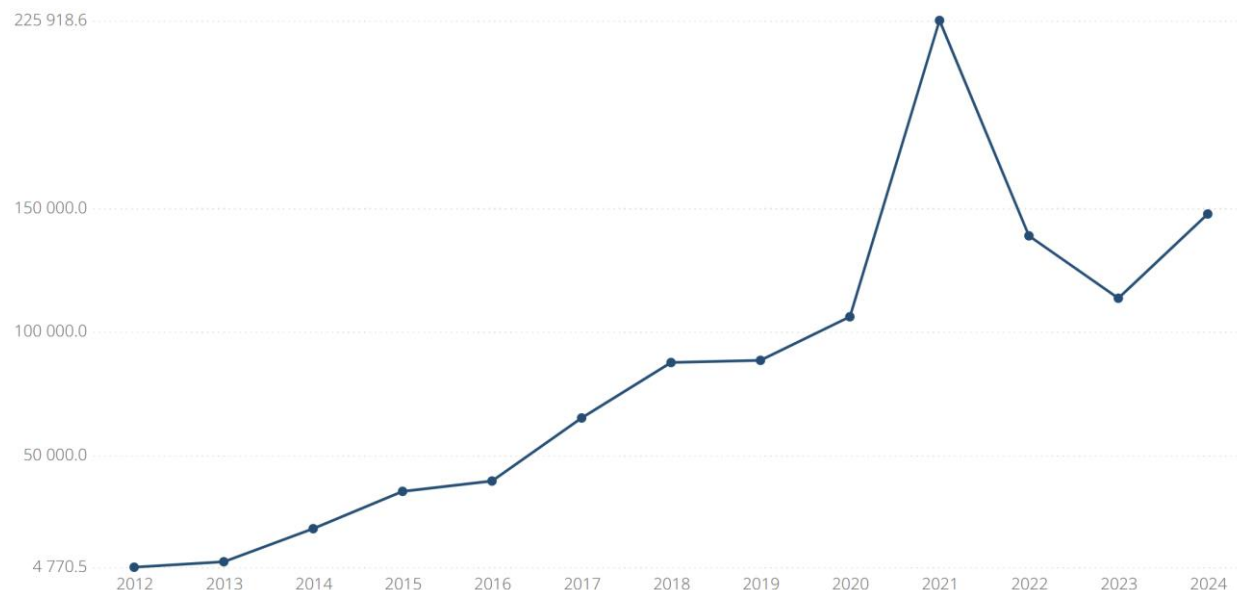
Note: A methodological note with more information can be found at https://oecd.ai/p/methodology. The surge in investments in 2021 was in part due to an increase in "healthcare, drugs and biotechnology" AI investments during the COVID-19 pandemic. A significant spike that year was also seen in "Mobility and autonomous vehicles".
Source: OECD.AI (2025), visualisations powered by JSI using data from Preqin, last updated 3 June 2025 accessed on 16 June 2025, www.oecd.ai.

Although some AI experts predict significant economic gains from AI, OECD (2024[24]) research indicates more tepid growth, estimating annual productivity growth due to AI ranging between 0.25-0.6 percentage points over the next 10 years in the most AI-ready countries. Research shows AI improves individual worker productivity (OECD, 2023[25]; Bengio et al., 2025[26]), but evidence connecting this to broader organisational and economic gains is weak. This is, in part, because some tasks cannot yet be conducted by AI and not all organisations or workers are ready to adopt it. Some evidence suggests firms adopting AI are more productive and grow faster than those that do not (Calvino and Fontanelli, 2023[27]; Hampole et al., 2025[28]), but this should not be interpreted as causality. For now, limitations persist. According to US Census Bureau statistics, only 5-6% of US businesses use AI to produce goods and services and only 7% plan to adopt AI in the coming months (Williams, 2025[29]). In a more global survey, only an estimated 26% of companies have the capabilities needed to derive real value from AI, and only 4% have succeeded in generating significant value (BCG, 2024[30]).[5]

Beyond economic gains, AI's transformative potential to achieve positive societal outcomes is beginning to show signs of promise. However, its full impact has yet to be realised. For instance, AI in science has contributed to real progress in robotics, nuclear fusion, drug discovery, antibody generation and protein folding (OECD, 2023[31]). Despite these early successes, many uses remain localised or experimental, and systemic change on a global scale is still forthcoming. AI's contribution to science is just beginning, and in some areas, the technology may have achieved less than anticipated. For example, some found AI contributed little to research during the COVID-19 pandemic (OECD, 2023[31]). So far, AI has mostly contributed to breakthroughs in a narrow set of natural and physical sciences. Similar transformations in other disciplines, such as social sciences, have progressed less despite high expectations (Manning, Zhu and Horton, 2024[32]). As such, while AI's societal benefits are emerging, the full scope of its transformative potential is still unfolding.

Using AI in government, a unique context

In addition to governing AI for society by setting the conditions and regulations for its trustworthy use, governments are striving to integrate the technology to better govern with AI. Similar to the private sector, the use of AI in government promises tremendous benefits while posing a number of risks and challenges. In fact, a Deloitte survey (2024[33]) of 2 770 senior leaders across 14 countries found public sector leaders were twice as likely as industry leaders to foresee AI-driven transformation in their organisations in the near term, but they felt more cautious and were less optimistic that it would result in productivity gains. Yet the subject has only recently become a focus of mainstream public management literature and many governments (Mergel et al., 2023[34]; Mellouli, Janssen and Ojo, 2024[35]). This is due to a combination of factors, including recent technological breakthroughs resulting in AI applications that are more practical and effective for government use (see Box 1.1); government access to vast amount of data that can be used as inputs for AI systems; and ongoing fiscal pressures that make AI attractive as a way to streamline operations and reduce costs. As a result, government use of AI trails that of the private sector.

While some lessons learned and success factors can be derived from industry efforts (Santos et al., 2024[36]), the purpose of and context within government are unique and present a number of specific opportunities and challenges. In addition, the field of AI is complex, progressing rapidly and has a steep learning curve for public servants and policymakers. If successful, however, the application of AI in government promises to significantly impact the wider economy and society by enhancing the quality and outcomes of public services, policies and government operations (Berglind, Fadia and Isherwood, 2022[37]).

Governments have huge influence over and impact in people's lives, bringing with it a duty of care for the public good — one that goes above that of companies (OECD, 2023[38]; Santiso, 2023[39]). Thus, they have a special responsibility to deploy AI in a way that minimises harm and prioritises the well-being of individuals and communities. This is especially the case when deploying AI in sensitive policy domains such as law enforcement, immigration control, welfare benefits and fraud prevention (OECD, 2024[13]).

Governments also operate with a unique mandate: they serve the public interest and are funded by public resources. As such, their actions — particularly those involving data and digital technologies — need to be guided by principles that uphold democratic values, individual rights and the rule of law. Unlike private entities, which may prioritise efficiency or profit, governments are expected to act transparently and with due regard for the public good to a greater extent than companies.

Key opportunity areas and benefits for AI in government

Opportunities for governments as developers and users of AI include the potential to transform service delivery, policymaking, internal operations and oversight. This is a pivotal moment for governments worldwide. Grappling with the rapid advances in AI technologies, they are trying to seize the opportunities provided by AI to innovate and modernise public administration, while managing and mitigating the associated risks, discussed below, and implementation challenges, discussed in Chapter 3.

Embracing AI in government unlocks new possibilities. Through years of research on the topic and working with governments around the world, the OECD (2024[13]) has identified three concrete opportunity areas for government use of AI:

- **Productivity** with more efficient internal operations and more effective policy design, decision-making and service delivery. For example, using predictive AI systems for more effective policy planning, automating processes for more accelerated service delivery, and boosting performance by allowing civil servants to focus less on mundane tasks and more on mission-critical activities.
- **Responsiveness** of public policies and service through enhanced design and delivery approaches that better meet evolving needs of citizens and specific communities, as well as through improved

civic participation mechanisms. This includes offering more personalised public services more proactively.

- **Accountability** by enhancing capacity for oversight and transparency, for instance through real-time monitoring. This shift may boost overall public satisfaction and enhance the perception of government as competent, fair and responsive, thereby strengthening public trust in government's capacity for innovation and transformation.

Table 1.1 illustrates how various AI tasks can feed into government activities, thus supporting these opportunity areas.

Table 1.1. Understanding the use of AI in government

AI tasks	Government activity	Opportunity area
- Recognition - Event detection - Forecasting - Personalisation - Interaction support - Goal-driven optimisation - Content generation - Reasoning with knowledge structures	Internal operations	Productivity (efficiency and effectiveness)
	Policymaking	Responsiveness
	Service delivery	
	Internal and external oversight	Accountability

Note: The AI tasks column is adapted from the "AI System Tasks" of the OECD Framework for the Classification of AI Systems (2022[40]).
Source: (OECD, 2024[13]).

Key benefits for AI in government

To guide investment decisions, it is crucial for public servants, especially decision-makers in leadership positions, to understand the benefits AI can offer. A study by the European Commission (EC) (2024[41]), which surveyed 576 public managers across seven countries, found that AI's perceived benefits significantly influence its adoption. AI has the potential to enhance decision-making at various stages of the policy cycle (Figure 1.3).[6] The sections below outline key benefits of using AI in government. These benefits are not mutually exclusive and indeed are highly complementary with some overlap in four concepts: automated, streamlined and tailored processes and services; better decision-making, sense-making and forecasting; enhanced accountability and anomaly detection; and unlocking opportunities for external stakeholders through AI as a good for all. It is important to note, however, that the use of AI in government can also pose risks. Several of these risks can be the converse of or undermine the potential benefits. The next section provides a dedicated discussion on risks.

Figure 1.3. AI at each stage of the policy cycle
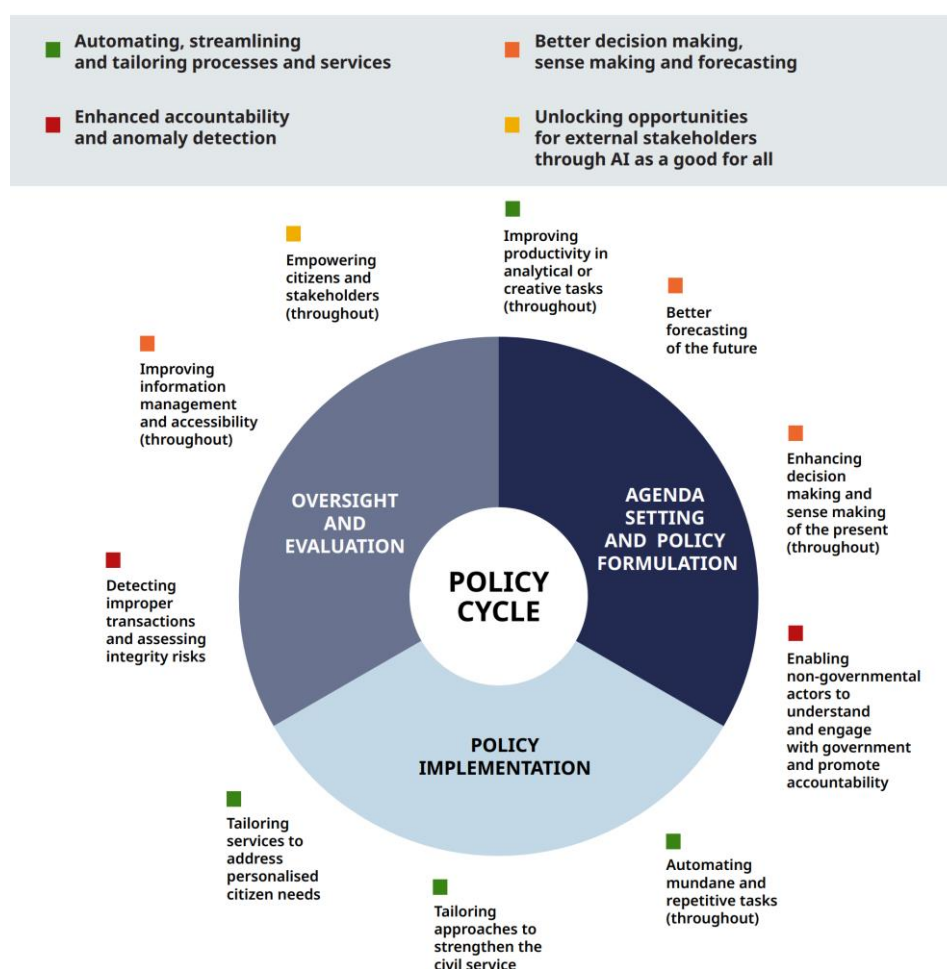
Source: (Pencheva, Esteve and Mikhaylov, 2018[42]), adapted to align with OECD terminology below.

Automated, streamlined and tailored processes and services

AI-enabled automation can help in directly automating existing processes and services, or contribute to the full re-imagining of how governments work both in internal operations as well as in public-facing services. In leveraging vast data assets, governments can also use AI to develop tailored services precisely crafted for specific individuals and groups. These benefits not only make government more efficient, effective and response, but they also can improve job quality and well-being for public servants by enabling them to spend more time on more valuable and meaningful work. This has been shown to improve workers' well-being. Nearly two-thirds of workers surveyed by the OECD (2023[25]) reported that AI improved their enjoyment of work, with studies showing this can enhance workers' well-being (Brougham and Haar, 2017[43]; Xu, Xue and Zhao, 2023[44]). However, as discussed below, some AI uses can reduce job quality and potentially lead to public service workforce displacement.

Automating mundane and repetitive tasks

Governments can use AI to the enhance efficiency of their internal operations and service delivery activities, reducing the time public officials invest in monotonous tasks (OECD, 2024[13])**.** Typically, these internal operations are repetitive and do not require extensive analytical thinking or human judgment. By automating these tasks, AI can streamline workflows, reduce errors, optimise resource allocation and free

up human resources for more complex, judgment-intensive activities. Ultimately, this leads to a more efficient delivery of higher quality public services (OECD/UNESCO, 2024[12]).

Repetitive and time-consuming tasks include:

- **Data entry**: manually inputting data into various systems and databases.
- **Payroll processing**: calculating and processing employee salaries and benefits.
- **Basic customer inquiries**: addressing routine inquiries and providing information to the public.
- **Information verification**: checking and verifying the authenticity of documents.
- **Form processing**: handling and processing various application forms.
- **Email and correspondence**: sorting, responding to and managing official emails and correspondence.

Governments can use of a variety of AI systems for these types of tasks, ranging from simplistic rules-based systems to more advanced ML systems, such as LLM-enabled chatbots. These systems have extensive capabilities ranging from handling simple and routine inquiries (both with citizens as well as with public servants) to generating entirely new tailored content and optimising resource allocation (Lorenz, Perset and Berryhill, 2023[17]; Sapci and Sapci, 2019[45]).

As an example, Argentina is automating repetitive tasks and expediting case processing in justice administration through its AI-driven Prometea system (see Chapter 5, Box 5.62). As discussed below, however, AI-enabled automation can pose risks — in areas such as justice administration — that governments need to consider. In the case of Prometea, for example, Argentina seeks to limit these risks through human control of how the AI system and its outputs are used (Corvalán and Le Fevre Cervini, 2020[46]).

Improving productivity in analytical or creative tasks

AI further contributes to productivity through the discovery of new ideas and through new efficient and effective means of conducting work (Jones, 2022[47]). One of AI's most notable benefits is its ability to help government offices handle and manage the analysis and synthesis of extensive documentation (OECD/UNESCO, 2024[12]). While a variety of AI-enabled tools may be useful, LLMs in particular can serve as powerful assistants for government officials in this regard, aiding in tasks such as research, content summarisation and synthesis (Berglind, Fadia and Isherwood, 2022[37]). Research conducted with knowledge professionals in the private sector showed that AI use can improve the performance of individual and teams, and also break down functional siloes (Dell'Acqua et al., 2025[48]). Governments can use AI to reduce civil servants' workload and to improve access to information for both citizens and public servants.

Some relevant uses include:

- **Processing and categorising textual information:** AI tools can quickly and accurately analyse extensive texts and unstructured documents, highlighting key points and summarising information. This enhances efficiency in departments dealing with vast amounts of information, such as legal affairs and administrative processes (OECD/UNESCO, 2024[12]). Consequently, workflows are sped up and the risk of human error is reduced, leading to more accurate and reliable outcomes in internal operations and in service delivery activities.
- **Drafting documents and legal texts**: AI systems can generate preliminary drafts of various types of documents by using templates and existing legislation. This process helps ensure adherence to standards while saving time and resources. Additionally, AI can cross-reference new drafts with current laws, identifying potential conflicts and minimising human errors. For report drafting, AI tools can offer automatic suggestions for clearer and more concise structures. Furthermore, it can

improve communication of extensive reports by summarising them into shorter formats for dissemination to decision-makers or the public.

- **Making sense of unstructured inputs:** AI can analyse and synthesise large amounts of information from participatory processes, public service feedback and consultations, turning them into actionable recommendations. It can identify recurring topics, cluster opinions, detect outliers, perform sentiment analysis and rank policy options based on preferences. This use can help identifying emerging issues, better considering stakeholder concerns and addressing potential policy impacts.

Studies show that generative AI systems reduce the time people spend conducting tasks while improving quality output. They also show that these tools have a greater impact on the productivity of lower-skilled workers, such as junior employees, allowing them to catch up with their more senior colleagues (Noy and Zhang, 2023[49]; Peng et al., 2023[50]). This can produce and equalising effect and drive productivity gains by helping workers conduct many tasks typically handled by subject matter experts (OECD, 2023[25]). However, some research also indicates that AI could contribute to further divides between higher-skilled and lower-skilled workers, and that while existing AI systems can increase worker productivity, they still cannot perform many tasks that humans can (The Economist, 2025[51]; Dell'Acqua et al., 2023[52]; Bengio et al., 2025[26]). More research is needed on this topic, including AI's specific advantages and potential drawbacks for public servants.

As an example, the UK tax authority is using AI to draft job descriptions and to analyse and evaluate the qualifications of job applications to speed up hiring (Box 5.20). As discussed below, using AI can pose risks if not done in a trustworthy manner.

AI can also play a role as a catalyst for creativity and innovation among public servants and how they design and implement internal processes and public policies and services. Generative AI, for instance, can support the exploration of policy alternatives, scenario simulations, legislative drafting and service prototyping, fostering a more imaginative and experimental public administration. For example, the UK's Government Communication Service is developing an AI-powered conversational tool to generate draft texts, plans and strategic ideas, integrating communication guidelines and audience insights to ensure high-quality, compliant outputs (Box 5.39). Designed as a collaborative assistant, it boosts creativity, reduces routine workload, and is being gradually rolled out after successful piloting and iterative AI-driven refinement.

Tailoring services to address personalised citizen needs

AI can help governments to better understand people's needs and behaviours, and facilitate the delivery of targeted and personalised information and services at an individual level (Huang and Rust, 2021[53]; Flavián and Casaló, 2021[54]; OECD, 2020[3]). This can include developing individualised citizen profiles, generating and delivering tailored information, and shaping service offerings based on unique needs (UN, 2022[55]). By improving responsiveness, such efforts can make services more efficient, effective and citizen-centred, resulting in higher satisfaction, better outcomes and a more agile and proactive approach to meeting public needs.

This enables a better response to the needs of user sub-groups, including vulnerable and disadvantaged groups, who may have specific and context-dependent needs (Giest, 2017[56]). The use of AI tools for personalisation has broadened into several public service sectors, frequently associated with life-event-based service delivery, such as services offered proactively for a child's birth, new educational pursuits or marriage (Kopponen et al., 2024[57]).

Governments can also tap into AI's capacity to analyse vast behavioural datasets for a deeper understanding of individual heterogeneity, incorporating cognitive and contextual factors — such as timing, location and personal preferences — into service design. This enables more adaptive and equitable

service interventions that align with citizens' unique circumstances while safeguarding autonomy and informed decision-making (Mills, Costa and Sunstein, 2023[58]).

In addition to enhancing services themselves, AI can help optimise communications on the availability and thus, the uptake of services. Using existing administrative data, AI can simplify processes by pre-filling (or eliminating) forms with known information and tailoring questions to the individual, reducing the time and effort required to complete bureaucratic tasks. This is already being done in a limited way with administrative data and (human) programmed algorithms in the area of social programmes (OECD, 2024[59]). This targeted approach can help ensure that citizens receive the information they need efficiently, and that interactions with public services are streamlined and user-friendly.

Examples of this benefit in action include a wide variety of chatbots and virtual assistants that can respond to unique queries from citizens with tailored information. For example, Singapore's tax authority offers a public-facing chatbot that can provide information and services tailored to individual needs (Box 5.4). As another example, social protection systems are using AI for proactive outreach to promote service uptake among those who quality (Box 5.49).

It is important to note, however, that such services often require a great deal of data collection and analysis to determine individual characteristics and needs. Governments should undertake such data collections and AI use in a trustworthy manner. Otherwise, as discussed below, these efforts run the risk of infringing upon the free exercise of human rights, such as through privacy infringements.

Tailoring approaches to strengthen the civil service

AI can also strengthen the public service through more effective and inclusive hiring processes and personalised programmes for continuous development. With regards to human resource management (HRM), for instance, AI can help governments optimise hiring decisions by helping to find the best candidates for the job and improve inclusiveness by controlling for human biases.

AI can also empower public servants; it can contribute to learning development, enhance knowledge creation and optimise learning platforms for upskilling public servants. This includes crafting skills development strategies, designing and running tailored training courses and implementing tools to improve information access. By doing so, AI can help ensure that public servants are equipped with the latest knowledge and skills necessary to meet the evolving demands of their roles, fostering a more responsive public administration.

Some relevant AI applications include:

- **Developing learning material for civil servants:** AI can create learning content (such as modules and course materials) from source documents and integrate diverse information into effective resources. It can also design, structure and deploy online learning courses. AI-driven tools can continually update and refine these resources as new information becomes available, enhancing the ongoing education and skill development of the government workforce.

- **Personalising material and learning routes for civil servants:** AI can tailor educational content and learning pathways to meet the specific needs, preferences and progress of each civil servant, ensuring a more effective, engaging and adaptive learning experience. By equipping civil servants with the right skills more efficiently, this personalisation also enhances government responsiveness.

- **Identifying and cataloguing learning resources:** AI tools can identify, describe and catalogue multiple learning resources, making them easily searchable and shareable, thereby simplifying resource discovery and identifying relevant content for learners. For example, integrating AI into digital platforms can enhance organisation, cataloguing and search functions.

As an example, the Australian Public Service Commission (APSC) has run a six-week pilot project to see if AI can design, structure and deploy an online learning course on digital skills for leadership (see Box 5.22). For further relevant discussion on this topic, see Chapter 5, section on "Civil service reform".

Better decision-making, sense-making and forecasting

AI experts have identified better decision-making, sense-making and forecasting as the most important AI benefits overall, and they recommend stronger actions and further investments by governments in achieving these areas (OECD, 2024[14]). As discussed below, governments should pursue these benefits while being cautious to avoid over-reliance on AI, as flaws in systems can be difficult to identify, and overly deferring human judgement to machines could allow the systemic propagation of errors and cause real-world harm.

Enhancing decision-making and sense-making of the present

By providing actionable, data-driven insights, AI can help governments improve the effectiveness and efficiency of targeting actions, allocate resources and identify policy problems and solutions. Governments can therefore respond more effectively to emerging issues, can ensure informed policy development, increase their overall responsiveness and accountability, and ultimately, promote societal well-being.

Some specific ways in which AI can be beneficial throughout the policy cycle (Figure 1.3) include:

- **Agenda setting and policy formulation**: AI can play a pivotal role bringing issues to the attention of policymakers and in the agenda setting process by framing social problems in ways that make them more responsive to actual social needs. For example, AI enables governments to monitor and make sense of emerging topics in real time from vast and representative datasets, enhancing the accuracy and speed of agenda-setting (Valle-Cruz et al., 2020[60]; Kolkman, 2020[61]). By detecting potential challenges more accurately and quickly, AI facilitates faster policy responses before issues escalate (OECD/UNESCO, 2024[12]; Höchtl, Parycek and Schöllhammer, 2015[62]). Through policy formulation, AI can influence the decision-making process by bringing important data and information about issues to the forefront (Valle-Cruz et al., 2020[60]). AI-enabled analysis can provide insights that estimate not only the likely impacts of policies, but also identify the target populations and make economic and social diagnosis, helping policymakers make informed decisions (Wirjo et al., 2022[63]; Ubaldi et al., 2019[9]). AI can also assist in devising policy alternatives, providing more in-depth ex-ante policy evaluation (Desouza and Jacob, 2014[64]).

- **Policy implementation**. As policies move to the implementation phase, AI-driven automation, rapid data processing and real-time analysis enhance the quality, speed and efficiency of policy implementation. AI analytics notably strengthen and expedite the acquisition of data and information, supporting continuous improvements. Real-time data analytics can facilitate large-scale enhancements, ultimately improving the delivery of services during policy implementation (Valle-Cruz et al., 2020[60]; OECD/UNESCO, 2024[12]).

- **Oversight and evaluation.** AI can monitor policy interventions in real time, providing better insights into the policy process, facilitating timely and accurate data assessments of policy interventions and enabling quick policy adjustments when needed (OECD, 2019[65]; OECD/UNESCO, 2024[12]).

Accessible data and increased computing power are providing AI with a competitive advantage over humans when it comes to decision-making in some cases (Green, 2022[66]). For instance, LLMs can support individual reasoning, and evidence shows real-world benefits from AI-assisted decision-making (Brynjolfsson, Danielle and Raymond, 2023[67]). AI systems can enhance decision =-making by mitigating reasoning mistakes and cognitive errors, helping humans filter out "noise" — the unwanted variability in human decision-making — and irrelevant influences that can lead to inconsistent and inaccurate decisions

(Du, 2023[68]). [7] The potential for AI systems to make data-driven decisions is leading to its adoption across a range of sectors, including within the public sector. AI can identify and address elements that distort human judgment across various applications in government (Mills, Costa and Sunstein, 2023[58]).

Additionally, noise may obscure key insights into human behaviour, which AI can uncover and quantify, contributing to more precise policymaking (Aonghusa and Michie, 2020[69]). Algorithms can significantly reduce noise by ensuring consistency in outcomes regardless of contextual factors, such as mood or time of day. By uncovering previously undetected or obscured behavioural patterns, AI can allow policymakers to better understand systemic trends and decision-making inconsistencies (Ludwig and Mullainathan, 2022[70]). While eliminating noise does not address all mistakes or errors, it enhances reliability, reducing arbitrary disparities in decisions across government functions like justice administration and public service delivery (Sunstein, 2023[71]).

As AI becomes more embedded in public administration, understanding the psychological and cognitive processes that shape human interactions with these technologies is increasingly important. Behavioural public administration can provide further valuable insights into the challenges that may arise, offering strategies to mitigate them and improve governance and decision-making (Alon-Barkat and Busuioc, 2024[72]). [8] An example of this is governments' use of Polis, an open-source civic engagement tool for understanding citizen views as well as areas of consensus and disagreement on public policy matters (Box 5.36). In the field of public financial management, Korea developed dBrain+, an information system that leverages AI to analyse real-time economic, fiscal and financial data to optimise risk assessment and decision-making in public finance (Box 5.8).

Better forecasting of the future

AI systems can process large volumes of data from multiple sources, including unstructured data, and identify complex patterns and weak signals — early signs of potential emerging changes, threats and opportunities — that are not easily detectable with existing methods. This can enhance the accuracy and timeliness of predictions and can be highly beneficial to other strategic foresight activities (Fitkov-Norris and Kocheva, 2025[73]). AI predictive analytics and forecasting in government involves using algorithms to anticipate future trends and risks. It can be widely applied across various domains, such as forecasting macroeconomic and fiscal outcomes (e.g. "nowcasting" GDP, as discussed in Chapter 5, section on "AI in public financial management").

By providing accurate and timely forward-looking insights, AI enhances decision-making, resource allocation and the overall effectiveness of government operations. Some relevant AI applications include:

- **Predicting future service needs**: AI has important potential in generating predictive analytics and forecasting, allowing public services to anticipate needs and be proactive. AI can help forecast future service needs, optimise resource allocation and enhance responsiveness across various policy domains by enabling the analysis of historical data and trends.

- **Regulatory forecasting**: regulators can use AI to uncover emerging trends and shifts in various industries to proactively plan regulatory responses. By continuously monitoring and analysing data from multiple sources, such as market reports, social media and news articles, AI can identify new developments and technological advancements that may impact regulatory landscapes.

- **Disaster risk management:** AI can also help forecast natural disasters by analysing historical data and current trends. For example, AI systems can analyse satellite imagery and other data to predict the likelihood of natural disasters like wildfires and earthquakes, allowing for proactive measures to minimise damage and enhance public safety (Sun, Bocchini and Davison, 2020[74]; Gupta and Roy, 2024[75]). AI systems can offer early warnings and useful information to help governments respond promptly to such events and mitigate their effects.

- **Anticipating corruption and fraud risks**: predictive AI systems are helping integrity actors to prioritise cases for further human examination. Although the transfer of these approaches from research to governments is still limited, it is growing steadily. For instance, AI systems can be used to prioritise risky cases and streamline auditing processes. They can also support anti-corruption policy targeting by providing early warning systems that predict public corruption based on data such as economic and political factors. Predictive techniques are also key to several AI-enabled government accountability and oversight activities, such as risk-based fraud detection, as further discussed below.

As an example of predicting future public service needs, Portugal's PrevOcupAI system aims to predict work-related illnesses and connected risks factors in the public administration to minimise disruptions (see discussion in Chapter 5, section on "AI in public service design and delivery").

Improving information management and accessibility

Robust, quality data is an underlying pre-requisite for enjoying AI's benefits. AI can help to maximise the quality and utility of data, as well as the ability of humans and machines to process and analyse it (Jarrahi et al., 2023[76]). For instance, AI systems enable new forms of data collection, including automatically detecting and identifying items in images, audio recordings or video. The capabilities and prevalence of AI-enabled sensing devices have progressed rapidly, allowing automatic speech transcription, motion detection, live image recognition and a wide range of tasks that previously required human labour (Zhang, Wang and Lee, 2023[77]; OECD, 2023[31]). AI can also improve how information is stored, disseminated and applied. This is particularly visible with the integration of AI into knowledge management systems (Sanzogni, Guzman and Busch, 2017[78]).

Improving how information is managed internally within governments can also help governments provide open information and data to the public. The use of AI can help minimise errors in data management, such as by reducing manual effort, and when used in combination with Privacy-Enhancing Technologies (PETs), can help enhance the privacy and protection of sensitive personal data and information (OECD, 2025[79]). This, in turn, enables the wider publication and access to open data. For instance, in justice administration, an AI-powered anonymisation engine can automatically identify and protect personal data within court decisions (Box 5.64), preparing them for public release as part of an open data initiative. The OECD (2023[80]) is also exploring the use of PETs, which are digital solutions that allow information to be collected, processed, analysed and shared while protecting data confidentiality and privacy.[9]

Internal government virtual assistants for civil servants provide a good example, with France's Albert and the UK's Caddy providing a wealth of historical and cross-government information at public servants' fingertips to inform decisions and help respond to citizen inquiries (Box 5.46).

Enhanced accountability and anomaly detection

One of the most longstanding uses of AI in government is to detect existing anomalies, such as fraud, or forecast future integrity risks, thus enhancing accountability and integrity in public programmes. Extra care in this use of AI should be taken to avoid possibly damaging outcomes in results.

Detecting improper transactions and assessing integrity risks

Such activities are common in a variety of government functions, including public procurement, tax administration and public financial management. Fraud and improper payments in government programmes can be significant sources of financial leakage. For instance, in the United States alone, the federal government loses an estimated USD 233-521 billion annually to fraud (US GAO, 2024[81]).

ML algorithms are particularly effective at pattern recognition, as they enable the analysis of large datasets and detect data outliers, hidden relationships (e.g. indicating collusion) and other anomalies that require further human investigation. Without the capacity to analyse data at scale or the capability to identify hidden patterns, such irregularities may go otherwise unnoticed without AI. This use of AI can enhance the ability of government organisations to maintain integrity and accountability. For example, France's tax authority uses AI to analyse arial photography and identify undeclared properties (Box 5.1).

Governments can use AI to better identify, evaluate, predict and respond to potential integrity risks, enabling better management, mitigation and timely interventions. For example, in regulatory compliance, inspectors are increasingly using AI to assess the risk posed by private operators. This improves the targeting of inspections, protection of public interests and efficient use of resources (OECD, 2018[82]; 2021[83]). AI assists inspectors by detecting patterns indicative of potential non-compliance, allowing for more accurate risk assessment. This use of AI not only streamlines inspections but also enhances the overall effectiveness of regulatory frameworks by ensuring that resources are directed where they are most needed.

Enabling non-governmental actors to understand and engage with government and promote accountability

Governments can use AI to be more transparent and to enable new forms and channels of interaction between citizens, civil society organisations (CSOs) and public institutions. In fact, AI experts suggest that empowering citizens, CSOs and social partners (e.g. trade unions) is one of the 10 most important benefits that AI can yield. This is underpinned by government transparency and use of AI (OECD, 2024[14]). They are:

- offering AI-enabled tools leveraging open government data (OGD) directly to citizens to help them navigate and make sense of government processes and actions
- enabling third-party oversight and scrutiny of government operations by CSOs and other non-governmental actors
- providing engagement opportunities and channels where the public can provide feedback and raise potential issues about government performance and decisions.

If executed well, this use of AI has the potential to promote accountability and public integrity, strengthen policymaking and increase citizen trust in government. This benefit — including real-world examples and discussion on the steps governments need to take to achieve it — are discussed in-depth in Chapter 5 in the section on "AI in civic participation and open government".

Unlocking opportunities for external stakeholders through AI as a good for all

A final benefit to the use of AI in government is unlocking new opportunities for external stakeholders, such as businesses and citizens, by providing them with access to government-developed AI systems. This benefit of AI may not be as directly observable in direct government activities as the other benefits. But it has the potential to improve public governance through increased trust in government, a more informed and skilled citizenry, or even economic growth. Countries have put in place OGD programmes not only to promote transparency and accountability, but also to promote entrepreneurship, economic growth and the creation of value that may not have necessarily been understood or foreseen in their onset (OECD, 2018[84]). For instance, Landsat satellite image data freely released by the US Geological Service since 2008 now results in over USD 25 billion annually in economic value (USGS, 2024[85]). When the OGD movement began in the late 2000s and early 2010s (Chignard, 2013[86]), it was not fully realised that the data made available would serve as vast resources for training AI systems a decade later.

AI's nature is distinctly different from that of data in that it is not a natural byproduct of an array of existing functions and activities, and it cannot serve as a raw input for other processes. (For example, data has been compared to fuel, electricity or drinking water for AI). Yet governments' use of AI has the potential to generate public good by empowering stakeholders to enhance their capabilities and access to information to derive new value.

Unlike other applications — where AI facilitates interactions between government and citizens, or improves information provision and accessibility — certain specific uses of AI in public governance can augment the capabilities of these actors, enabling them to achieve their missions and objectives more effectively. For example, the government of India has made AI systems available to farmers to help them ensure crop health and mitigate pest-related challenges (Jeevanandam, 2024[87]). Another example is Germany's FAIR Forward – Artificial Intelligence for All development initiative, which supports open creation and responsible usage of AI systems on areas such as agriculture, climate protection and citizen participation (OECD, 2023[88]). Such approaches may be particularly useful in markets not yet served by, or with little appeal for, private sector solutions. Governments may have the resources to invest in under-explored fields and can take the first steps to build out new markets, taking risks where others may not be ready or willing to.

The potential for positive external opportunities would be amplified further if governments were to provide access to enablers for AI, such as digital infrastructure in the form of computing power (Ho, 2023[89]) (see Chapter 4, section on "Building out digital infrastructure"). In some cases, these enablers may already exist for government use, only requiring some adjustments and scaling to make them available to a broader audience. In other cases, enablers targeting external actors could be developed and supplied. This can democratise AI's potential value (OECD, 2024[14]).

Non-governmental entities can also participate in the creation and deployment of AI tools, playing leading roles in fostering technology-enabled participation (OECD, 2025[90]).

To seize the benefits of AI, its risks need to be managed

The global adoption of AI in all sectors raises questions about trust, fairness, privacy, safety and accountability, among others. Considering these issues and managing AI's risks can have an impact on its adoption and the realisation of its benefits (Tse and Karimov, 2022[91]). AI poses hundreds of risks,[10] and experts identify some of the most important for society as (OECD, 2024[15]; [14]; [92]; 2022[93]):

- Possible adverse outcomes for some groups or individuals if AI systems are underpinned by inadequate or skewed data.
- AI systems lacking sufficient transparency and explainability erode accountability.
- Certain AI uses raise data protection, privacy and surveillance concerns.
- AI can facilitate increasingly sophisticated cyber threats.
- Minor to serious AI incidents and disasters could occur in critical systems.
- AI could contribute to labour market disruptions.
- AI computational power requires very significant energy use.

These are thorny issues that are being considered by governments, companies, CSOs and other relevant stakeholders. Efforts such as the adoption of the European Union (EU) AI Act (2024[94]) and a variety of national-level governance initiatives help to illustrate how governments are taking an active role in governing AI for society.[11]

Risks specific to AI use in government

Perhaps all AI risks could implicate governments in some way, such as necessitating governance processes or mitigation measures. Yet a narrower subset of risks is most relevant for policymakers and public servants as they pursue the strategic and responsible use of AI in government. As mentioned above, governments have an outsize influence on people's lives, thus their use of AI can have a greater impact on the public in both positive and negative ways. Accountability expectations are, therefore, higher for government use of AI. This can be seen explicitly in the EU AI Act, which classifies many public sector AI use cases as "high-risk", and some as "unacceptable risk" and therefore banned in the EU (Box 1.2). The United States takes a different approach, considering some use cases as "high-impact", and thereby requiring certain risk management practices (Box 1.3). Korea's Basic Act on the Development of AI and the Establishment of Foundation for Trustworthiness ("AI Basic Act") (2024[95]), which will take effect in January 2026 and applies to organisations developing or using AI in the Korean market, similarly designates some uses as "high-impact", thereby requiring enhanced measures for ensuring AI safety and reliability. The AI Basic Act also includes a separate designation for generative AI that includes specific transparency and disclosure requirements.

Box 1.2. The EU AI Act's risk levels as related to AI in government

The European Union (EU) AI Act is a regulation on AI that entered into force in August 2024. The regulation establishes obligations for AI based on its potential risks and level of impact. The Act identifies four different levels of risks that are relevant for governments' use of AI.

- **Unacceptable risk**. AI uses under this category are prohibited. Examples include predictive policing, real-time remote biometric identification (including facial recognition) in public spaces for law enforcement, social scoring and assessing the risk of an individual committing criminal offenses. Law enforcement and justice administration are among the functions of government most concerned by this category. However, 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, many government 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 for 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, setting up technical documentation to demonstrate compliance and mandatory fundamental rights impact assessment.

- **Limited risk**. These uses might include systems intended to communicate with individuals, such as chatbots, as well as systems that generate content such as text and images. Transparency obligations require developers and deployers to 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.

Source: (EU, 2024[94]).

Box 1.3. The concept of "high-impact" AI in United States (US) policy

In the United States, the policy *M-25-21 Accelerating Federal Use of AI through Innovation, Governance and Public Trust* establishes a dualistic approach for AI use cases in government: either the AI use case is considered "high-impact", or it is not.

High-impact AI is AI with an output that serves as a principal basis for decisions or actions with legal, material, binding or significant effect on:

- an individual or entity's civil rights, civil liberties or privacy; access to education, housing, insurance, credit, employment, critical government resources or services; or other programmes
- human health and safety
- critical infrastructure or public safety
- strategic assets or resources, including high-value property and information marked as sensitive or classified by the federal government.

Federal agencies are responsible for conducting reviews on their AI use cases and determining the applicability of the high-impact definition. The policy provides examples of 15 categories of AI use cases that are presumed to be high-impact.

Agencies must implement **minimum risk management practices** for high-impact AI use cases. The minimum risk management practices include the following: conducting pre-deployment testing and an AI impact assessment; ongoing monitoring for performance and potential adverse impacts; ensuring adequate human training, assessment, human oversight, intervention and accountability; offering consistent remedies and appeals; and consulting and incorporating feedback from end users and the public. Agencies must have a plan to discontinue the use of any high-impact AI system that is not performing at an appropriate level in compliance with the policy, until actions are taken to achieve compliance.

Limited pilot programmes should follow minimum risk management practices where practicable. However, if the CAIO certifies and other criteria are met as detailed in the source document, pilot programmes are exempt from the minimum risk management practices. Agency CAIOs may also waive one or more minimum risk management practices under certain circumstances for a specific use case, though they must certify the ongoing validity of each determination and waiver annually, track them centrally and publicly release a summary of each.

Source: https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf.

As they seek to develop and use AI, governments face risks that include potential dangers and threats that could cause serious problems for individuals and society (Valle-Cruz, Garcia-Contreras and Gil-Garcia, 2023[96]), potentially undermining public trust, the legitimacy of government's AI use and even democratic values. To address these concerns, it is important to identify and manage these risks, consider how AI systems may impact citizens or marginalised populations differently, help ensure the equitable distribution of AI benefits and mitigate potential harm. The continuous consideration of potential risks is important because known risks can evolve and new risks can emerge, including ones previously considered to be outside the realm of possibility.

This report identifies five general types of risks for the use of AI in government, as presented below. Beyond grappling with these risks, governments also face a number of implementation challenges when seeking to develop and use AI. These implementation challenges are discussed in Chapter 3.

- **Ethical risks**: These include AI uses that undermine the free exercise of human rights and freedoms, including privacy, potentially infringing on human-centred values either deliberately or inadvertently. AI algorithms can introduce ethical risks from the digital realm to the physical world through biased algorithms and unethical behaviours like invasive surveillance. Key concerns include threats to trust, fairness, freedom, dignity, individual autonomy and labour rights.
- **Operational risks**: These include technical and operational failures that might affect data privacy, the quality of AI outcomes and internal government operations due to cybers threats, unintended consequences, hallucinations, systematic errors and overreliance on AI systems.
- **Exclusion risks**: These risks relate to gaps that arise when citizens without access to technology or digital literacy can be left behind and unable to benefit from AI advancements in public services.
- **Public resistance risks**: These include public resistance to government use of AI. This can be driven by distrust in government AI systems or processes, or by the spread of false or misleading information about how AI is implemented in public administrations and its potential impacts.
- **Risks of inaction**: Although often overlooked, this risk includes government delays in using AI to yield positive benefits. This can result in significant financial and non-financial costs — which could have otherwise been avoided with AI's successful adoption — and a widening gap between public sector and private sector capabilities.

Ethical risks

Inadequate or skewed data in AI systems

AI systems have the potential to perpetuate or generate adverse or harmful outcomes, stemming from incomplete or inadequate data, as well as how AI usage intersects with institutional and social practices that are human-centred or systemic in nature.

It is important to recognise that algorithms do not operate autonomously; they are shaped by human choices at every stage, from model selection and training data to fine-tuning and parameter adjustments. Since AI systems usually learn from human-generated data, they inevitably reflect existing social outlooks and behaviours. Moreover, in government, algorithms rarely make decisions independently — they typically serve as tools that inform and influence human decision-making rather than replace it. Even where policies aim for standardisation, successful implementation still heavily depends on local context and the "messy engagement of multiple players with diverse knowledge" (Davies, Nutley and Walter, 2008[97]). This interplay between AI and human judgment means that errors can persist not only in algorithmic outputs but also in how AI-generated recommendations are interpreted and applied. Understanding how decision-makers process and respond to AI-generated information and ensuring they have the skills to use AI in a trustworthy manner are therefore critical. On one side, cognitive subjectivity remains relevant in shaping human-AI interactions despite AI's perceived neutrality, while on the other, machines lack empathy. Therefore, it is important that public servants can use professional judgment and practical wisdom to ensure fairness by exercising judicious discretion.

AI systems are also highly sensitive to the quality of training data, making them susceptible to overfitting patterns,[12] learning spurious correlations and amplifying errors or skew embedded in human-generated datasets. While AI call help eliminate noise, as discussed as a benefit above, its pattern-seeking nature can also exacerbate it, particularly when systemic inconsistencies exist in the training data. Without rigorous oversight and mitigation, AI risks reinforcing distortions rather than delivering objective, reliable decisions, highlighting the critical need for careful data curation and validation, system testing, anticipatory

and retrospective impact monitoring and assessment, and algorithmic transparency in AI-driven decision-making (Shane, 2019[98]).

Because of its potential to have systemic impacts, insufficient or skewed data and algorithms can also exacerbate other types of risks, such as those discussed below.

Misuse or questionable use of AI, resulting in surveillance and privacy concerns

The misuse or questionable application of AI in government could result in harm to the free exercise of individual freedoms and rights. A prominent concern is the use of AI in delivering public security and safety services, where it enables efficient identification and tracking through biometric data and real-time monitoring. While these tools can be useful for law enforcement and crime prevention, they also raise concerns about data privacy, as well as surveillance and social control by public administrations. For instance, during the COVID-19 pandemic, AI was employed to track people's movements to ensure compliance with self-isolation mandates (Saheb, 2022[99]). Although this measure was intended to control the spread of the virus, it raised privacy concerns among the public (OECD, 2020[100]). Indeed, AI-driven surveillance in increasingly pervasive, with the Carnegie Endowment for International Peace (2022[101]) finding that 97 of 179 (54%) countries analysed are using AI technologies for public surveillance.[13] The index identified non-democratic states as a major driver of AI surveillance, including through selling products to other countries. Yet others, including liberal democracies, are also major users of AI surveillance (Saheb, 2022[99]). As seen in Box 1.2, the EU AI Act regulates when the use of AI for surveillance is admissible, and it identifies use cases that are an "unacceptable risk" and banned throughout the EU.

Similarly, personalised service delivery often requires algorithmic processing of data scattered across multiple public and sometimes private data sources (Nikiforova et al., 2023[102]). This raises privacy concerns and highlights the need for governments to pursue both efficient service delivery and the robust protection of individual privacy rights. For instance, to provide personalised social services, governments might aggregate data from healthcare records, educational achievements, employment history and even social media activity. While the goal is to offer tailored support and timely interventions, processing such vast amounts of personal information needs to be done carefully and deliberately, with controls in place to mitigate concerns of surveillance. Such guidance may be found in (OECD, 2017[103]) and (OECD, 1980[104]). Robust policies and infrastructure are needed to consider trade-offs between responsiveness, transparency and the protection of sensitive information (OECD, 2024[105]; [59]).

Another misuse case is social scoring in service delivery or policymaking, a practice where individuals are classified based on behaviour or personal traits. AI algorithms analyse data from sources like social media, financial transactions and public records to assign scores. These scores can impact access to services, loans and employment opportunities, leading to unfair treatment. Government use of such systems is banned in the EU as an unacceptable use (EU, 2024[94]).

In a recent survey of hundreds of experts across fields, 79% said that AI will have a negative impact on people's privacy by 2040, a concern shared by the general public (Rainie and Anderson, 2024[106]; Fazlioglu, 2024[107]). Governments will need to ensure their use of AI is trustworthy to allay these concerns as it related to AI in government.

The potential misuse of AI tools for citizen surveillance by authorities can lead to overreach and abuse of power. Continuous monitoring and data collection can create a climate of fear and mistrust, particularly among communities who may already feel disproportionately targeted by law enforcement (UN, 2024[108]). Governments can also use AI to strengthen political power, potentially facilitating wide-scale subjugation and authoritarianism (OECD, 2024[14]), especially by non-democratic governments or those that do not prioritise the protection of human rights. Some experts argue that AI could be — and in some instances already is being — used to track and monitor citizens and residents at scale, using algorithms and

behavioural analysis, identify and supress opposition, and perpetuate totalitarian regimes (OECD, 2022[109]; Tegmark, 2017[110]; Clarke and Whittlestone, 2022[111]; Byler, 2021[112]).

Finally, algorithmic manipulation, where AI systems and their results are altered to produce specific outcomes, is another potential misuse and unethical behaviour in the governmental use of AI (Valle-Cruz, Garcia-Contreras and Gil-Garcia, 2023[96]). This manipulation can stem from individual public servants, decision-makers or AI system developers intentionally altering system results to benefit or harm certain groups or individuals or push a specific agenda. AI's inherent complexity can additionally make it challenging to trace and understand algorithms' inner workings.

Lack of transparency and explainability

Systems based on deep learning are "black boxes", meaning that it is difficult to describe how they produce a given output. Such outputs are indirectly generated from deep learning training as engineers continually tweak parameters until the model scores highly on training objectives (Clarke and Whittlestone, 2022[111]). Even scientists working on advanced deep learning models do not understand the inner workings of their systems, and they find it hard to trace back these outputs and test the reliability of these systems through traditional methods (OECD, 2022[109]).

This makes it hard to detect and mitigate harmful outcomes and produces challenges in determining accountability when issues arise. As AI systems become increasingly integrated into functions of government, black box systems could make it difficult to explain the rationale for AI-assisted decisions to citizens. It can also exacerbate other AI risks. For instance, it can be more difficult to identify algorithmic bias and its root causes in opaque systems. Public servants could also increasingly have a false sense of trust in seemingly efficient yet flawed AI systems because such flaws may be unobservable (contributing to automation bias, discussed below) (OECD, 2024[14]; Russell, 2019[113]). This can erode government accountability and disempower the public by limiting their ability to make informed decisions or potentially making them subject to opaque, flawed AI-driven decisions (Lima et al., 2022[114]).

Operational risks

"Automation bias" – overreliance on AI

Many people perceive AI systems and their decisions to be neutral and impartial, leading users to accept results without scrutiny. Studies on AI-assisted decision-making have identified a tendency to overweight algorithmic recommendations, often assuming their prediction to be more reliable than human judgment — even when the AI system itself has limitations (Alon-Barkat and Busuioc, 2024[72]). This "automation bias" (the propensity for people to trust AI outputs because they appear rational and neutral) can lead to the application of AI decision-making to a growing number of societal challenges (Horowitz, 2023[115]; Alon-Barkat and Busuioc, 2022[116]) — perhaps to avoid difficult conversations and decisions about human approaches to these issues. Some experts assert that this habit is creating "blind faith" in technology, a problematic phenomenon that can reinforce existing systemic issues against certain groups or individuals and contribute to neglect of human suffering and erosion of empathy (Goldman, 2023[117]; Olson, 2023[118]).

So-called "automation bias" in government occurs when public organisations or civil servants rely too heavily on AI systems for decision-making or task execution. For example, if healthcare professionals rely too heavily on AI-automated suggestions without cross checking, they might miss critical information or make incorrect diagnoses. This excessive dependence can result in users failing to recognise mistakes, accepting incorrect AI outputs, and diminishing human oversight and judgment (Passi and Vorvoreanu, 2022[119]; Klingbeil, Grützner and Schrec, 2024[120]). AI could be systematically adopted without fully assessing accuracy and potential consequences, leading to reliance on AI systems and the propagation of compounding errors throughout entire systems (Valle-Cruz, Garcia-Contreras and Gil-Garcia, 2023[96]).

One contributor to such issues are hallucinations, which occur when generative AI systems make up facts in a credible way, often when a correct answer is not found in the training data. This can be harmful in contexts such as government decision-making, where they may lead to misguided decisions or actions (OECD, 2024[15]; Beltran, Ruiz Mondragon and Han, 2024[121]). For example, an AI system designed to answer queries from the public could provide erroneous details about public services or list services that do not exist. If not carefully mitigated, these unintended consequences and incorrect outputs can scale rapidly, affecting large populations or critical internal government decisions.

AI systems can also be prone to simple errors and malfunctions, which can lead to systematic deviations and imprecision in algorithmic outputs. If these systems are not carefully implemented or fail to be technically reliable, they risk diminishing public trust, including that of civil servants. Citizens may experience frequent frustrations due to incorrect processing of their requests or delays in service delivery. For example, automated systems that handle form submissions, customer service inquiries and payment requests need to operate with great accuracy; otherwise, mistakes or technical glitches can lead to a perception of incompetence and unreliability in public services. Technical problems or errors in chatbot interactions, such as incorrect responses, inability to understand queries or system outages, can further erode trust, making citizens feel that AI systems are unreliable and ineffective.

Risk aversion also plays a role in "automation bias". Civil servants may be afraid to take personal responsibility for decisions for fear of getting in trouble later. An example is when a human decides against the advice of an AI system and it turns out to be the wrong decision. In such cases, the civil servant will seem doubly responsible. As seen in Box 5.43, some government AI systems have even been designed to make public servants write justifications for review if they go against an AI system's advice, adding to the burden of their work and reinforcing incentives to follow what the system recommends.

Research also highlights the concept of selective adherence, where decision-makers are more likely to follow algorithmic recommendations when they align with their pre-existing beliefs or societal stereotypes (Alon-Barkat and Busuioc, 2022[116]). This can lead to distorted public decision-making, as civil servants may unconsciously use AI outputs to justify pre-existing biases rather than critically evaluating them.

Some research suggests that overreliance on AI may contribute to a decline in human cognitive abilities, reducing exploration, creativity and independent thought, as individuals become accustomed to AI-generated solutions. Studies suggest that frequent reliance on AI-driven decision support systems can lead to cognitive offloading, reducing individuals' engagement in critical thinking and independent problem-solving (Gerlich, 2025[122]). AI-driven decision-making also risks promoting behavioural homogenisation, as its outputs often reflect limited diversity of perspectives. This narrowing of perspectives could hinder adaptive thinking and reduce the capacity of societies and governments to navigate uncertainty and risk (Meng, 2024[123]). Although over-reliance on AI may pose risks, AI tools can also help human operators interpret and question complex AI decisions, discouraging overreliance (OECD, 2024[15]).

Reduced job quality for public servants

While AI can enhance public servants' job quality and well-being, as discussed above, some uses can have the opposite effect. The use of algorithmic management tools is increasing significantly, reaching an adoption rate of 90% in US firms and 79% in the EU (Milanez, Lemmens and Ruggiu, 2025[124]). While public sector-specific studies have not been conducted, tangible concerns have been raised about existing negative impacts of AI and algorithmic tools on job quality, including work intensification, increased stress, perceived reduction in fairness and workplace surveillance (OECD, 2023[25]). For instance, AI could make jobs less fulfilling and more stressful by incentivising new types of surveillance in the workplace, or new forms of hyper-efficient yet exhausting "digital Taylorism", in which work is subject to increased surveillance and regulation, including through algorithmic management (UC Berkeley, 2021[125]; EC, 2025[126]).[14] Further research has shown that such AI surveillance can harm mental health (APA, 2023[127]), and that AI task management can erode the autonomy and voice of workers, reducing human insights into how work is

managed (Gmyrek, Berg and Bescond, 2023[128]). However, if used well, such tools have also been shown to improve worker safety and well-being (e.g. by alerting workers about dangers and hazards, or identifying burnout) (EC, 2025[126]). Algorithmic management and its impacts are already being seen and studied with regard to their use in the private sector (OECD, 2023[25]). Research on algorithmic management in the public sector is scarce, although its use is expanding rapidly (EC, 2025[126]).

Privacy and data governance tensions

Developing and deploying AI systems poses privacy and data governance challenges throughout the AI lifecycle (OECD, 2024[129]; forthcoming[130]). In the AI training stage, many developers depend on publicly accessible sources for building AI training datasets, which may purposefully or inadvertently include personal data or information subject to intellectual property rights. However, just because data is accessible on the Internet does not automatically mean that it is free to be collected and used to train AI models. Further, people may have shared their personal data consenting to another use or uses, which do not necessarily include training AI models (ICO, 2023[131]). The collection of personal data for training AI systems, like any data processing activity, is subject to the commonly recognised privacy principles set forth in the OECD Recommendation Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data (OECD Privacy Guidelines) (1980[104]). These principles require that personal data be obtained through lawful and fair means, with the knowledge of the data subject, and that any further uses of the data are not incompatible with the original purposes.

Another important aspect to consider is the capacity of AI models to memorise personal data within their parameters during the training stage. As a result, LLMs behind text-based generative AI tools pose a particular risk of unauthorised access and use of third-party personal data without the knowledge of the individuals concerned (Brown et al., 2022[132]). Some research also shows that generative AI models are able to infer personal attributes of the data subject from text with high accuracy, yet at a low cost (Staab et al., 2023[133]). This raises privacy concerns not only because these inferences can reveal personal information or personal characteristics, especially when such traits were not intended to be shared.

During the deployment stage, AI systems can also be in tension with individuals' rights to access, correct, and where necessary, delete their personal data (also known as the "Individual Participation Principle" in the OECD Privacy Guidelines). For example, fulfilling individuals' rights to have their data deleted or corrected can be technically complex and resource-intensive, as it might require identifying specific data points related to an individual within unstructured datasets or in some cases re-training the AI model. Additionally, research conducted prior to the widespread use of generative AI models suggests that, in certain cases, it is possible to reconstruct or deanonymize original training data by analysing the behaviour of a model that includes that data (Salem et al., 2018[134]). To address such tensions, promoting further international cooperation between data privacy and AI communities can contribute to harmonised data practices with AI development and use. For example, OECD's Expert Group on AI, Data, and Privacy[15] is exploring policy responses on data governance and privacy in the context of AI, involving experts from multiple sectors and disciplines around the world.

Cyber threats

AI systems incorporated into government processes can be vulnerable to cyber threats, which can lead to data breaches, privacy violations and loss of functionality. The extensive collection and analysis of personal data for AI applications can result in loss, alteration or unauthorised disclosure of this data, infringing on individual privacy rights (Beltran, Ruiz Mondragon and Han, 2024[121]). Unauthorised access and data breaches can compromise personal data and operational integrity, which could result in identity theft, financial fraud and other privacy violations, undermining public trust in government institutions and resulting in legal consequences. Additionally, malicious cyber actors could manipulate AI systems, altering their outputs and causing erroneous decisions or actions (Brundage et al., 2018[135]; Gopireddy, 2024[136]).

Cyber risks may originate from external bad actors, or from insider threats within government (Eshelby, 2025[137]). Overall, cybersecurity can be seen as a horizontal function of government in itself, and indeed represents one of the areas of greatest AI adoption in government for enhancing security of government IT systems (Mariani, Kishnani and Alibage, 2025[138]). However, this function is highly specialised and not included in the primary scope of this report, and it is not a topic of thorough analysis and discussion. However, the OECD has an ongoing workstream on digital security that has conducted relevant work.[16]

Exclusion risks

Exacerbating digital divides

The risk of omitting people when using AI in government is closely linked to digital divides (UN, 2024[108]), the gap between those who can access and use information and communication technologies and those who cannot. This is particularly evident in public service delivery and policymaking activities, especially in the use of AI for predictive analytics, forecasting and service personalisation.

Digital divides risks hindering access to the benefits and public value that AI offers. This is especially the case among populations lacking the necessary infrastructure and digital literacy to engage with AI-driven public services (OECD, 2024[14]; ITU, 2023[139]). AI in government can offer advantages — like tailored information, faster response times and enhanced service delivery — but may be inaccessible to those without internet access or digital skills. Governments' shift to digitalisation can intensify the barriers to digital government services among certain segments of the population. Research in Canada found that users in rural locations, women and girls and those in low-income households were negatively affected by the push towards digital-first services due to the COVID-19 pandemic (Singh and Chobotaru, 2022[140]). In Norway, the digitalisation and automation of the system for awarding child benefits made it possible for most recipients to receive the benefit automatically. At the same time, it resulted in the need for other recipients to apply manually, a burden that disproportionally affected low-income segments (Larsson, 2021[141]).

Data divides are another form of exclusion caused by AI that is linked to unrepresentative data, a consequence of existing digital divides. Individuals without internet access are often absent from the data used to develop algorithms. Because most data are predominantly collected through the internet and online interactions, the widespread use of AI may exclude these individuals, as algorithms used to inform policymaking lack representative data. These gaps make it difficult for governments to respond adequately to the needs of all citizens, resulting in insufficient or inadequate data sets in AI-driven decision-making processes and service delivery. For instance, data divides limit the potential for AI benefits, such as personalised AI services, leaving them only useful and accurate for data-rich populations (UNESCO, 2019[142]; Perry and Turner Lee, 2019[143]; Dieterle, Dede and Walker, 2022[144]).

Another form of digital divide involves underrepresentation of languages (Röttger et al., 2024[145]; Peixoto, Canuto and Jordan, 2024[146]). Training datasets tend to overrepresent widely used languages, as seen in Figure 1.4. This imbalance can lead to AI systems that fail to serve non-dominant language groups effectively. Language preservation efforts, such as Estonia's Donate a Speech, reduce the current limitations of speech technology, which favour the most widely recognised languages, and enhance service delivery (OECD, 2023[38]). Additional considerations and examples are discussed in Chapter 4, section on "Creating a strong data foundation".

Figure 1.4. More than half (59%) of open-source AI training datasets are in English

Percentage breakdown of languages for open-source AI training datasets on Hugging Face
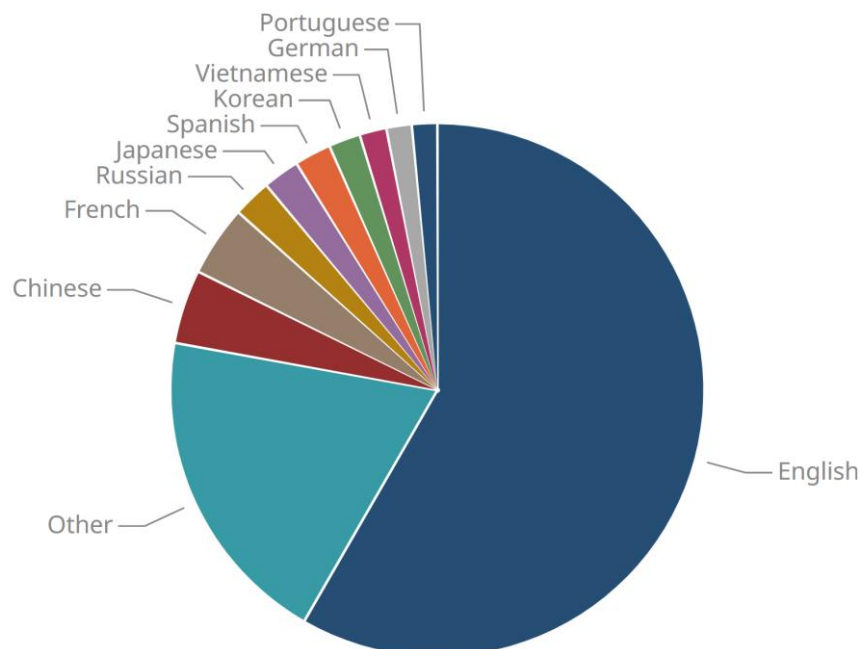
Note: This chart represents the language distribution of all datasets. Multilingual and translation datasets on Hugging Face contain more than one language and are thus double counted. More methodological information available at: https://oecd.ai/huggingface.
Source: OECD.AI (2024), visualisations powered by JSI using data from Hugging Face. Last updated on 5 June 2025, (accessed 16 June 2025).

Public service workforce displacement

The OECD (2023[25]) found that while AI is capable of automating non-routine tasks, its future impacts on labour are ambiguous; they depend on the balance between the displacement of human labour by AI, the increase in labour demand due to greater productivity brought about by AI and the creation of new jobs caused by AI adoption. AI can enhance government productivity by automating tasks, and it also has the potential to reduce the need for human labour, resulting in the need to re-skill public servants to take on more meaningful tasks (Peixoto, Canuto and Jordan, 2024[146]).

At the same time, AI deployment is driving a growing demand for AI-related skills across the economy, and the public sector is no exception. As AI-related roles increase, hiring and retention efforts for non-AI or traditional positions may decline, highlighting the need for a workforce skilled in AI and related technologies to meet evolving demands (Acemoglu et al., 2022[147]).

AI technologies impact a wide range of occupations and sectors, affecting workers of all skill levels and influencing labour markets (OECD, 2023[25]). While some government workers may adapt to AI and even see their work enhanced, others, such as older and low-skilled workers conducting tasks that are easy to automate, face significant risks. This suggests that AI's benefits are not evenly shared among public servants (Milanez, 2023[148]). For instance, AI-powered chatbots are now commonly used for service delivery and citizen-centred communication to answer basic questions and provide information, which reduces the demand for government human customer service representatives (Acemoglu, 2024[149]).

For many years, concerns about job automation focused on low-skilled labour, but high-cognitive tasks are increasingly being assisted by AI, which can impact civil servants such as policy analysts. Generative AI can produce meaningful text, conduct data analysis and even propose policy strategies to address complex

challenges. Many civil servants will need to be trained to collaborate effectively with AI and focus on higher level strategic thinking and decision-making.

Public resistance risks

Citizens may selectively accept AI-informed outputs, potentially contributing to errors

Understanding human-AI interactions is increasingly seen as a key challenge for public administration, with significant implications for trust and legitimacy. As governments integrate AI into decision-making, public acceptance of these systems becomes critical. Research suggests algorithmic decisions, despite their promise of neutrality, are not always perceived as fair or legitimate, particularly when they contradict individuals' expectations or lack transparency (Alon-Barkat and Busuioc, 2022[116]).

Citizens process algorithmic decisions through cognitive perceptions and prior beliefs; they often selectively accept AI-generated recommendations that align with their expectations while resisting those to the contrary (Alon-Barkat and Busuioc, 2022[116]). This selective adherence can reinforce errors in government decision-making, as individuals may be more likely to trust algorithmic predictions when they confirm pre-existing stereotypes or prior knowledge.

Lack of public empowerment and understanding about how government uses AI

There is often a lack of knowledge and understanding among the public regarding AI in general, and more specifically, how governments use it (Arnesen et al., 2024[150]). This can lead to misconceptions and fears about its capabilities and whether governments are using it in a trustworthy manner. This could potentially result in outcomes such as rumours spread online about government's use of AI in policymaking or in delivering services such as loans, access to justice or social benefits. This can occur because AI's complexity is often misunderstood, leading to inaccurate assumptions about how decisions are made, the fairness of outcomes and the ability to hold these systems accountable; or because of the limited transparency on how public authorities are using AI. This limited transparency or gap in understanding makes the public more susceptible to scepticism regarding governments' actions and intentions, resulting in resistance to AI-driven solutions in public services (Valle-Cruz, Garcia-Contreras and Gil-Garcia, 2023[96]).

AI decisions can be **perceived as overly rigid or unaccountable**, particularly in the absence of human oversight or avenues for redress. Research suggests that when citizens feel disempowered in interactions with automated systems — such as being unable to challenge incorrect AI-based decisions — they experience higher psychological and compliance costs (Alon-Barkat and Busuioc, 2024[72]).

Conversely, higher levels of public empowerment and transparency on use, knowledge and understanding of AI are associated with greater trust in the government's ability to use AI responsibly (Lahusen, Maggetti and Slavkovik, 2024[151]; KPMG, 2025[152]; Alessandro et al., 2021[153]). A well informed citizenry is more likely to recognise the safeguards and ethical considerations implemented by the government, fostering confidence in its competence and integrity in deploying AI technologies.

AI misuse and scandals can undermine trust and contribute to public resistance

The use of AI in government has resulted in several high-profile scandals and cases of real-world harm. This underscores the high reputational costs of AI misuse, with issues happening years ago still resident in public discourse. High-profile failure in one AI system can erode confidence in a broader array of government use cases (Longoni, Cian and Kyung, 2022[154]). Public backlash has even led to people withdrawing from data sharing or hindered the use of existing tools (Ada Lovelace Institute, 2025[155]).

Public backlash is particularly likely when AI errors result in visible harm, such as wrongful denial of benefits or unfair treatment based on skewed predictions. High-profile failures can reinforce public scepticism and fuel concerns AI-driven decisions may undermine procedural justice and democratic accountability (Alon-Barkat and Busuioc, 2022[116]). Data from the OECD AI Incidents Monitor (AIM) shows growth in AI incidents and hazards reported by reputable media sources in recent years (Figure 1.5).[17] As of April 2025, 3 816 of the 14 981 incidents listed (25%) were related to "government, security and defence", illustrating that governments need to ensure they mitigate risks in order to secure citizen trust.

Figure 1.5. AI incidents have generally trended upwards since late 2022
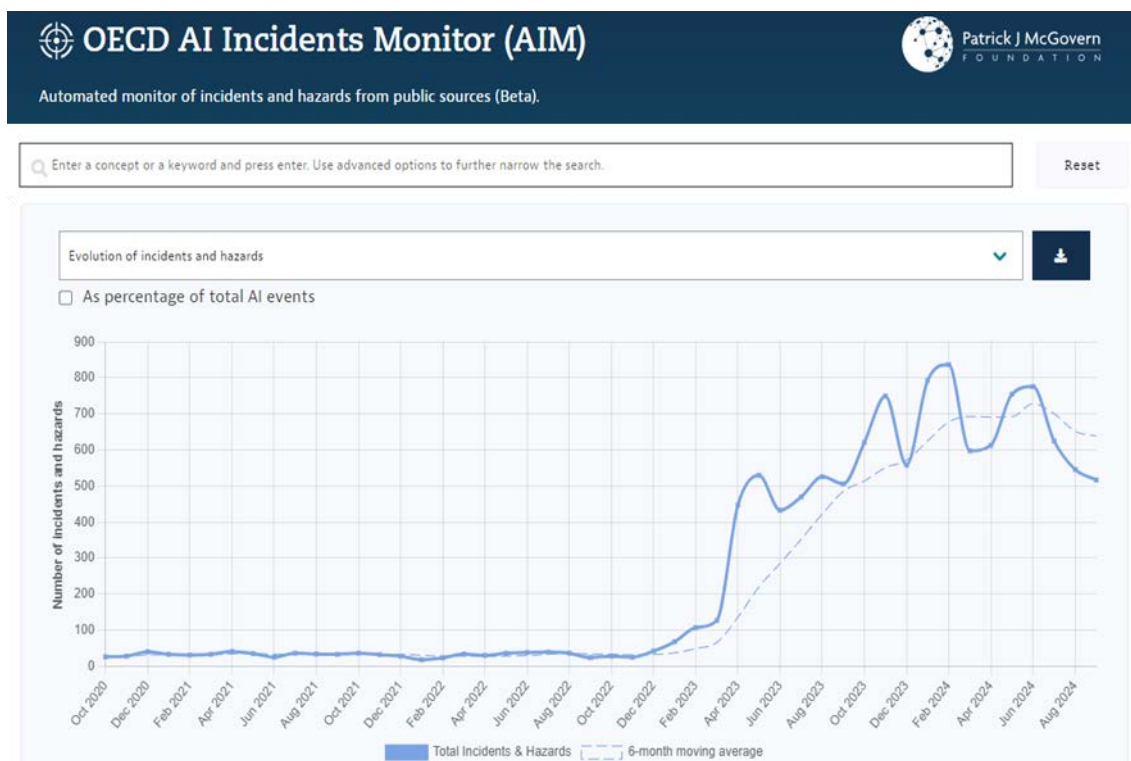
Note: An overview of the methodology can be found at https://oecd.ai/incidents-methodology.
Source: OECD AI Incidents Monitor (AIM) – https://oecd.ai/incidents.

Citizens can also exhibit algorithmic aversion, often resisting algorithmic decision-making due to a preference for personal agency and control, even when AI systems outperform humans. This reluctance is reinforced by a greater tolerance for human errors compared to algorithmic mistakes; people tend to lose trust in AI after observing a single failure, whereas they are more forgiving of similar errors made by humans. Algorithmic aversion can be more common in some functions (e.g. public safety) than in others (e.g. general management), which can contribute to different challenges and differing levels of maturity across government functions (see Chapter 3) (Zehnle, Hildebrand and Valenzuela, 2025[156]). Providing users with insight into the reasoning behind an AI's recommendation or offering limited control to modify its output can significantly improve acceptance, increasing the likelihood of users adopting AI-driven advice (Sunstein, 2023[71]). Beyond contributing to public resistance risks, algorithmic aversion on the part of public servants can hinder governments ability to harness the benefits of AI, as discussed in Chapter 3.

Risks of inaction

The most discussed risks of AI involve the implications of its deployment and adoption. However, a less discussed risk involves delays in leveraging AI to yield real-world positive benefits, including in government and public services. Stanford University's *One Hundred Year Study on Artificial Intelligence (AI100)* (2014[21]) noted that "numerous advances in AI can reduce costs, introduce new efficiencies and raise the quality of life... However, the methods have not come into wide use. The sluggish translation of these technologies into the world translates into unnecessary deaths and costs. There is an urgent need to better understand how we can more quickly translate valuable existing AI competencies and advances into real-world practice." While this study is over a decade old, the conclusion remains the same, especially in government. Beyond missed opportunities, the risk of inaction on AI serves to widen the gap between public sector and private sector capacity (Pahlka, 2024[157]). This not only means governments could fall behind on their ability to use AI but also in their ability to regulate the technology. AI experts suggest that one of the most critical AI risks is the inability of governance mechanisms and institutions to keep up with rapid AI evolutions (OECD, 2024[14]).

Some research highlights that negative hype and fear around AI can contribute to this risk (Laplante et al., 2020[158]). Such research notes that other limitations may include a lack of suitable data, confusion around privacy issues and complexities, and dealing with outdated legacy IT systems. Experts have also attributed the issue of "analysis paralysis" — the fear of getting AI wrong — as a possibility that could paralyse the implementation of even low-risk efforts, foregoing significant benefits (OECD.AI, 2023[159]). This has proven true in government; the latest cross-cutting OECD (2024[13]) work on AI in government found that governments need to better promote and enable the positive aspects of using AI, rather than focusing so disproportionately on the prevention of negative ones. This focus on risks might deter the deployment of high-benefit, low-risk uses of AI to improve public policies and services.

Realising a positive future for AI in government

As discussed above, governments are seeking to use AI to increase productivity, with more efficient internal operations and more effective policies and services; responsiveness, through tailored approaches that meet the evolving needs of citizens and businesses; and accountability by enhancing their capacity for oversight. However, there is tremendous untapped potential for governments to use today's AI technologies and prepare to unlock the opportunities presented by tomorrow's AI. Even so, a vision of a future where governments successfully develop and adopt trustworthy AI is beginning to emerge.

> *See AI not as an opportunity to automate the public sector but to reimagine it. We welcome a long-term vision for public service transformation where AI follows rather than leads, one that is grounded in public and professional legitimacy. Public sector leaders should see the rollout of AI as an opportunity to reimagine the state, rather than focusing solely on immediate efficiency gains or automating the status quo. AI should be viewed as a catalyst for fundamental service redesign, placing the citizen at the centre of public service delivery. – Ada Lovelace Institute (2025[155])*

This vision shines through in Chapter 5's discussion of AI in core functions of government. By pursuing AI as part of their digital journeys, governments can transform, rather than just optimise, how they achieve their missions, deliver public value and promote societal well-being.

Future capabilities and uses of AI may present benefits and changes that are currently impossible or even inconceivable. This is also true of its potential risks. The contents of Chapter 5 represent what is currently known about AI in government, and by extension, what can be imagined about the future. Governments and the OECD need to remain vigilant in evaluating how evolving AI technologies and applications may affect public institutions, civil servants, and society at large — ensuring continuous assessment and adaptation in service of the public good.

Through analysis and synthesis of the government functions in Chapter 5 and other research, and for further discussion of AI in government, the OECD has conducted cross-cutting research into the current trends of AI use across government functions and has identified early lessons from these use cases. This research and its findings are discussed in the following chapter.

References

Acemoglu, D. (2024), "The Simple Macroeconomics of AI", *NBER, Working Paper 32487*, http://www.nber.org/papers/w32487. [149]

Acemoglu, D. et al. (2022), "Artificial Intelligence and Jobs: Evidence from Online Vacancies", *Journal of Labor Economics*, Vol. 40/S1, pp. S293-S340, https://doi.org/10.1086/718327. [147]

Ada Lovelace Institute (2025), *Learn fast and build things: Lessons from six years of studying AI in the public sector*, Ada Lovelace Institute, https://www.adalovelaceinstitute.org/policy-briefing/public-sector-ai/. [155]

Alessandro, M. et al. (2021), "Transparency and Trust in Government. Evidence from a Survey Experiment", *World Development*, Vol. 138, p. 105223, https://doi.org/10.1016/j.worlddev.2020.105223. [153]

Alon-Barkat, S. and M. Busuioc (2024), "Public administration meets artificial intelligence: Towards a meaningful behavioral research agenda on algorithmic decision-making in government", *Journal of Behavioral Public Administration*, Vol. 7, https://doi.org/10.30636/jbpa.71.261. [72]

Alon-Barkat, S. and M. Busuioc (2022), "Human–AI Interactions in Public Sector Decision Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice", *Journal of Public Administration Research and Theory*, Vol. 33/1, pp. 153-169, https://doi.org/10.1093/jopart/muac007. [116]

Aonghusa, P. and S. Michie (2020), "Artificial intelligence and behavioral science through the looking glass: Challenges for real-world application.", *Annals of Behavioural Medicine*, pp. 942-947, https://doi.org/10.1093/abm/kaaa095. [69]

APA (2023), *Worries about artificial intelligence, surveillance at work may be connected to poor mental health*, https://www.apa.org/news/press/releases/2023/09/artificial-intelligence-poor-mental-health. [127]

Arnesen, S. et al. (2024), "Knowledge and support for AI in the public sector: a deliberative poll experiment", *AI & SOCIETY*, https://doi.org/10.1007/s00146-024-02104-w. [150]

Austin, T. et al. (2024), *A snapshot of how public sector leaders feel about generative AI*, https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-adoption-in-public-sector.html. [33]

BCG (2024), *Where's the Value in AI?*, https://www.bcg.com/publications/2024/wheres-value-in-ai. [30]

Beltran, M., M. Ruiz Mondragon and S. Han (2024), "Comparative Analysis of Generative AI Risks in the Public Sector", *Proceedings of the 25th Annual International Conference on Digital Government Research*, pp. 610-617, https://doi.org/10.1145/3657054.3657125. [121]

Bengio, Y. et al. (2025), *International AI Safety Report*, DSIT 2025/001, 2025, https://www.gov.uk/government/publications/international-ai-safety-report-2025. [26]

Berglind, N., A. Fadia and T. Isherwood (2022), *The potential value of AI—and how governments could look to capture it*, https://www.mckinsey.com/industries/public-sector/our-insights/the-potential-value-of-ai-and-how-governments-could-look-to-capture-it (accessed on July 2024). [37]

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. [8]

Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, European Commission, https://op.europa.eu/s/z4XY. [22]

Brougham, D. and J. Haar (2017), "Smart Technology, Artificial Intelligence, Robotics, and Algorithms (STARA): Employees' perceptions of our future workplace", *Journal of Management & Organization*, Vol. 24/2, pp. 239-257, https://doi.org/10.1017/jmo.2016.55. [43]

Brown, H. et al. (2022), "What Does it Mean for a Language Model to Preserve Privacy?", *2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 2280-2292, https://doi.org/10.1145/3531146.3534642. [132]

Brundage, M. et al. (2018), *The Malicious Use of Artificial Intelligence: Forecasting, Precention, and Mitigation*, https://arxiv.org/abs/1802.07228. [135]

Brynjolfsson, E., L. Danielle and L. Raymond (2023), *Generative AI at Work*, National Bureau of Economic Research, https://doi.org/10.3386/w31161. [67]

Byler, D. (2021), *In the Camps: China's High-Tech Penal Colony*, Columbia Global Reports, https://www.jstor.org/stable/j.ctv2dzzqqm. [112]

Calvino, F. and L. Fontanelli (2023), "A portrait of AI adopters across countries: Firm characteristics, assets' complementarities and productivity", *OECD Science, Technology and Industry Working Papers*, No. 2023/02, OECD Publishing, Paris, https://doi.org/10.1787/0fb79bb9-en. [27]

Chignard, S. (2013), *A brief history of Open Data*, https://www.paristechreview.com/2013/03/29/brief-history-open-data. [86]

Clarke, S. and J. Whittlestone (2022), *A Survey of the Potential Long-term Impacts of AI - How AI Could Lead to Long-term Changes in Science, Cooperation, Power, Epistemics and Values*, https://dl.acm.org/doi/abs/10.1145/3514094.3534131. [111]

Cognitus, A. (2024), *9 Agentic AI Examples: Real-World Use Cases and Applications*, https://integrail.ai/blog/agentic-ai-examples. [18]

Corvalán, J. and E. Le Fevre Cervini (2020), *Prometea experience. Using AI to optimize public institutions*, https://ceridap.eu/prometea-experience-using-ai-to-optimize-public-institutions. [46]

DataHeroes (2023), *Noise in Machine Learning*, https://dataheroes.ai/glossary/noise-in-machine-learning. [160]

Davies, H., S. Nutley and I. Walter (2008), "Why 'knowledge transfer' is misconceived for applied social research", *Journal of Health Services Research & Policy*, Vol. 13/3, pp. 188-190, https://doi.org/10.1258/jhsrp.2008.008055. [97]

Dell'Acqua, F. et al. (2025), *The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise*, Elsevier BV, https://doi.org/10.2139/ssrn.5188231. [48]

Dell'Acqua, F. et al. (2023), "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4573321. [52]

Desouza, K. and B. Jacob (2014), "Big Data in the Public Sector: Lessons for Practitioners and Scholars", *Administration & Society*, Vol. 49/7, pp. 1043-1064, https://doi.org/10.1177/0095399714555751. [64]

Dieterle, E., C. Dede and M. Walker (2022), "The cyclical ethical effects of using artificial intelligence in education", *AI & SOCIETY*, Vol. 39/2, pp. 633-643, https://doi.org/10.1007/s00146-022-01497-w. [144]

Du, M. (2023), "Machine vs. human, who makes a better judgment on innovation? Take GPT-4 for example", *Frontiers in Artificial Intelligence*, Vol. 6, https://doi.org/10.3389/frai.2023.1206516. [68]

EC (2025), *Study exploring the context, challenges, opportunities, and trends in algorithmic management*, European Commission, https://employment-social-affairs.ec.europa.eu/study-exploring-context-challenges-opportunities-and-trends-algorithmic-management_en. [126]

EC (2024), *What factors influence perceived artificial intelligence adoption by public managers*, https://publications.jrc.ec.europa.eu/repository/handle/JRC138684. [41]

Eshelby, L. (2025), *Addressing insider threats in the public sector*, https://www.openaccessgovernment.org/addressing-insider-threats-in-the-public-sector/187801/. [137]

EU (2024), *Regulation (EU) 2024/1689 laying down harmonised rules on artificial*, European Union, https://eur-lex.europa.eu/eli/reg/2024/1689/oj. [94]

Fazlioglu, M. (2024), *Consumer Perspectives of Privacy and Artificial Intelligence*, https://iapp.org/resources/article/consumer-perspectives-of-privacy-and-ai/. [107]

Feldstein, S. (2022), *AI & Big Data Global Surveillance Index (2022 updated)*, https://doi.org/10.17632/gjhf5y4xjp.4. [101]

Filippucci, F., P. Gal and M. Schief (2024), "Miracle or Myth? Assessing the macroeconomic productivity gains from Artificial Intelligence", *OECD Artificial Intelligence Papers*, No. 29, OECD Publishing, Paris, https://doi.org/10.1787/b524a072-en. [24]

Fitkov-Norris, E. and N. Kocheva (2025), "Leveraging AI for strategic foresight: Unveiling future horizons", in *Improving and Enhancing Scenario Planning*, Edward Elgar Publishing, https://doi.org/10.4337/9781035310586.00023. [73]

Flavián, C. and L. Casaló (2021), "Artificial intelligence in services: current trends, benefits and challenges", *The Service Industries Journal*, Vol. 41/13-14, pp. 853–859, https://doi.org/10.1080/02642069.2021.1989177. [54]

Gartner (2024), *Gartner 2024 Hype Cycle for Emerging Technologies Highlights Developer Productivity, Total Experience, AI and Security*, https://www.gartner.com/en/newsroom/press-releases/2024-08-21-gartner-2024-hype-cycle-for-emerging-technologies-highlights-developer-productivity-total-experience-ai-and-security. [23]

Gerlich, M. (2025), "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking", *Societies*, Vol. 15/1, p. 6, https://doi.org/10.3390/soc15010006. [122]

Giest, S. (2017), "Big data for policymaking: fad or fasttrack?", *Policy Sciences*, Vol. 50/3, pp. 367-382, https://doi.org/10.1007/s11077-017-9293-1. [56]

Gmyrek, P., J. Berg and D. Bescond (2023), *Generative AI and Jobs: A global analysis of potential effects on job quantity and quality*, ILO, https://doi.org/10.54394/FHEM8239. [128]

Goldman, S. (2023), *OpenAI has grand 'plans' for AGI. Here's another way to read its manifesto*, https://venturebeat.com/ai/openai-has-grand-plans-for-agi-heres-another-way-to-read-its-manifesto-the-ai-beat/. [117]

Gopireddy, R. (2024), *Securing AI Systems: Protecting Against Adversarial Attacks and Data Poisoning*, https://jsaer.com/download/vol-11-iss-5-2024/JSAER2024-11-5-276-281.pdf. [136]

Government of Korea (2024), *A New Chapter in the Age of AI: Basic Act on AI Passed at the National Assembly's Plenary Session*, https://www.msit.go.kr/eng/bbs/view.do?bbsSeqNo=42&mId=4&mPid=2&nttSeqNo=1071. [95]

Green, B. (2022), "The flaws of policies requiring human oversight of government algorithms", *Computer Law & Security Review*, Vol. 45, p. 105681, https://doi.org/10.1016/j.clsr.2022.105681. [66]

Grzegorzek, J. (2024), *Digital Taylorism: The Use of Data to Monitor Employees*, https://medium.com/%40JerryGrzegorzek/digital-taylorism-the-use-of-data-to-monitor-employees-582b331d970a. [161]

Gupta, T. and S. Roy (2024), "Applications of Artificial Intelligence in Disaster Management", *Proceedings of the 2024 10th International Conference on Computing and Artificial Intelligence*, pp. 313-318, https://doi.org/10.1145/3669754.3669802. [75]

Hampole, M. et al. (2025), *Artificial Intelligence and the Labor Market*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w33509. [28]

Höchtl, J., P. Parycek and R. Schöllhammer (2015), "Big data in the policy cycle: Policy decision making in the digital era", *Journal of Organizational Computing and Electronic Commerce*, Vol. 26/1-2, pp. 147-169, https://doi.org/10.1080/10919392.2015.1125187. [62]

Ho, D. (2023), *Opportunities and Risks of Artificial Intelligence in the Public Sector*, https://law.stanford.edu/2023/05/25/opportunities-and-risks-of-artificial-intelligence-in-the-public-sector/. [89]

Horowitz, M. (2023), *Bending the Automation Bias Curve: A Study of Human and AI-based Decision Making in National Security Contexts*, https://arxiv.org/abs/2306.16507. [115]

Horvitz, E. (2014), *Reflections and Framing: One-Hundred Year Study on Artificial Intelligence: Reflections and Framing*, https://ai100.stanford.edu/reflections-and-framing. [21]

Huang, M. and R. Rust (2021), "Engaged to a Robot? The Role of AI in Service", *Journal of Service Research*, Vol. 24/1, pp. 30–41, https://doi.org/10.1177/1094670520902266. [53]

ICO (2023), *Joint statement on data scraping and the protection of privacy*, https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2023/08/joint-statement-on-data-scraping-and-data-protection/. [131]

ITU (2023), *Measuring digital development: Facts and Figures 2023*, https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx (accessed on July 2024). [139]

Jarrahi, M. et al. (2023), "Artificial intelligence and knowledge management: A partnership between human and AI", *Business Horizons*, Vol. 66/1, pp. 87-99, https://doi.org/10.1016/j.bushor.2022.03.002. [76]

Jeevanandam, N. (2024), *AI in agriculture in 2025: Transforming Indian farms for a sustainable future*, https://indiaai.gov.in/article/ai-in-agriculture-in-2025-transforming-indian-farms-for-a-sustainable-future. [87]

Jones, C. (2022), "The Past and Future of Economic Growth: A Semi-Endogenous Perspective", *Annual Review of Economics*, Vol. 14/1, pp. 125-152, https://doi.org/10.1146/annurev-economics-080521-012458. [47]

Klingbeil, A., C. Grützner and P. Schrec (2024), *Trust and reliance on AI — An experimental study on the extent and costs of overreliance on AI*, https://doi.org/10.1016/j.chb.2024.108352 (accessed on September 2024). [120]

Kolkman, D. (2020), "The usefulness of algorithmic models in policy making", *Government Information Quarterly*, Vol. 37/3, p. 101488, https://doi.org/10.1016/j.giq.2020.101488. [61]

Kopponen, A. et al. (2024), "Personalised public services powered by AI: the citizen digital twin approach", in *Research Handbook on Public Management and Artificial Intelligence*, Edward Elgar Publishing, https://doi.org/10.4337/9781802207347.00020. [57]

KPMG (2025), *Trust in artificial intelligence: global insights 2025*, https://kpmg.com/au/en/home/insights/2025/04/trust-in-ai-global-insights-2025.html. [152]

Lahusen, C., M. Maggetti and M. Slavkovik (2024), "Trust, trustworthiness and AI governance", *Scientific Reports*, Vol. 14/1, https://doi.org/10.1038/s41598-024-71761-0. [151]

Laplante, P. et al. (2020), "Artificial Intelligence and Critical Systems: From Hype to Reality", *Computer*, Vol. 53/11, pp. 45-52, https://doi.org/10.1109/mc.2020.3006177. [158]

Larsson, K. (2021), "Digitization or equality: When government automation covers some, but not all citizens", *Government Information Quarterly*, Vol. 38/1, p. 101547, https://doi.org/10.1016/j.giq.2020.101547. [141]

Lima, G. et al. (2022), "The Conflict Between Explainable and Accountable Decision-Making Algorithms", *2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 2103-2113, https://doi.org/10.1145/3531146.3534628. [114]

Longoni, C., L. Cian and E. Kyung (2022), "Algorithmic Transference: People Overgeneralize Failures of AI in the Government", *Journal of Marketing Research*, Vol. 60/1, pp. 170-188, https://doi.org/10.1177/00222437221110139. [154]

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.

[17]

Ludwig, J. and S. Mullainathan (2022), "Algorithmic Behavioral Science: Machine Learning as a Tool for Scientific Discovery", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4164272.

[70]

Manning, B., K. Zhu and J. Horton (2024), *Automated Social Science: Language Models as Scientist and Subjects*, https://arxiv.org/abs/2404.11794.

[32]

Mariani, J., P. Kishnani and A. Alibage (2025), *Government's less trodden path to scaling generative AI*, https://www2.deloitte.com/us/en/insights/industry/public-sector/government-faces-challenges-with-generative-ai-adoption.html.

[138]

Mellouli, S., M. Janssen and A. Ojo (2024), "Introduction to the Issue on Artificial Intelligence in the Public Sector: Risks and Benefits of AI for Governments", *Digital Government: Research and Practice*, Vol. 5/1, pp. 1-6, https://doi.org/10.1145/3636550.

[35]

Meng, J. (2024), *AI emerges as the frontier in behavioral science*, https://doi.org/10.1073/pnas.2401336121.

[123]

Mergel, I. et al. (2023), "Implementing AI in the public sector", *Public Management Review*, pp. 1-14, https://doi.org/10.1080/14719037.2023.2231950.

[34]

Milanez, A. (2023), "The impact of AI on the workplace: Evidence from OECD case studies of AI implementation", *OECD Social, Employment and Migration Working Papers*, No. 289, OECD Publishing, Paris, https://doi.org/10.1787/2247ce58-en.

[148]

Milanez, A., A. Lemmens and C. Ruggiu (2025), "Algorithmic management in the workplace: New evidence from an OECD employer survey", *OECD Artificial Intelligence Papers*, No. 31, OECD Publishing, Paris, https://doi.org/10.1787/287c13c4-en.

[124]

Mills, S., S. Costa and C. Sunstein (2023), "AI, Behavioural Science, and Consumer Welfare", *J Consum Policy*, Vol. 46, pp. 387–400, https://doi.org/10.1007/s10603-023-09547-6.

[58]

Nikiforova, A. et al. (2023), "Towards High-Value Datasets Determination for Data-Driven Development: A Systematic Literature Review", in *Lecture Notes in Computer Science, Electronic Government*, Springer Nature Switzerland, Cham, https://doi.org/10.1007/978-3-031-41138-0_14.

[102]

NIST (2025), *Technical Blog: Strengthening AI Agent Hijacking Evaluations*, https://www.nist.gov/news-events/news/2025/01/technical-blog-strengthening-ai-agent-hijacking-evaluations.

[20]

Noy, S. and W. Zhang (2023), "Experimental evidence on the productivity effects of generative artificial intelligence", *Science*, Vol. 381/6654, pp. 187-192, https://doi.org/10.1126/science.adh2586.

[49]

OECD (2025), *How Innovation Ecosystems Foster Citizen Participation Using Emerging Technologies in Portugal, Spain and the Netherlands*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/2cb37a30-en.

[90]

OECD (2025), *Sharing trustworthy AI models with privacy-enhancing technologies*, OECD Publishing, https://doi.org/10.1787/a266160b-en.

[79]

OECD (2025), "Towards a common reporting framework for AI incidents", *OECD Artificial Intelligence Papers*, No. 34, OECD Publishing, Paris, https://doi.org/10.1787/f326d4ac-en. [162]

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. [5]

OECD (2024), "AI, data governance and privacy: Synergies and areas of international co-operation", *OECD Artificial Intelligence Papers*, No. 22, OECD Publishing, Paris, https://doi.org/10.1787/2476b1a4-en. [129]

OECD (2024), "Assessing potential future artificial intelligence risks, benefits and policy imperatives", *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, https://doi.org/10.1787/3f4e3dfb-en. [14]

OECD (2024), "Defining AI incidents and related terms", *OECD Artificial Intelligence Papers*, No. 16, OECD Publishing, Paris, https://doi.org/10.1787/d1a8d965-en. [163]

OECD (2024), "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en. [7]

OECD (2024), *Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity*, OECD Publishing, Paris, https://doi.org/10.1787/d909ff7a-en. [92]

OECD (2024), *Global Trends in Government Innovation 2024: Fostering Human-Centred Public Services*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/c1bc19c3-en. [105]

OECD (2024), "Governing with Artificial Intelligence: Are governments ready?", *OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, https://doi.org/10.1787/26324bc2-en. [13]

OECD (2024), *Modernising Access to Social Protection: Strategies, Technologies and Data Advances in OECD Countries*, OECD Publishing, Paris, https://doi.org/10.1787/af31746d-en. [59]

OECD (2024), *OECD Artificial Intelligence Review of Germany*, OECD Publishing, Paris, https://doi.org/10.1787/609808d6-en. [11]

OECD (2024), *OECD Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier*, OECD Publishing, Paris, https://doi.org/10.1787/a1689dc5-en. [15]

OECD (2024), *OECD Survey on Drivers of Trust in Public Institutions – 2024 Results: Building Trust in a Complex Policy Environment*, OECD Publishing, Paris, https://doi.org/10.1787/9a20554b-en. [4]

OECD (2024), *Recommendation of the Council on Artificial Intelligence*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. [6]

OECD (2023), *Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research*, OECD Publishing, Paris, https://doi.org/10.1787/a8d820bd-en. [31]

OECD (2023), "Emerging privacy-enhancing technologies: Current regulatory and policy approaches", *OECD Digital Economy Papers*, No. 351, OECD Publishing, Paris, https://doi.org/10.1787/bf121be4-en. [80]

OECD (2023), *Global Trends in Government Innovation 2023*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/0655b570-en. [38]

OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, https://doi.org/10.1787/08785bba-en. [25]

OECD (2023), "The state of implementation of the OECD AI Principles four years on"*, OECD Artificial Intelligence Papers*, No. 3, OECD Publishing, Paris, https://doi.org/10.1787/835641c9-en. [88]

OECD (2022), *Emerging future AI risks*, OECD, https://wp.oecd.ai/app/uploads/2023/03/OECD-Foresight-workshop-notes-1.pdf. [109]

OECD (2022), "Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint"*, OECD Digital Economy Papers*, No. 341, OECD Publishing, Paris, https://doi.org/10.1787/7babf571-en. [93]

OECD (2022), "OECD Framework for the Classification of AI systems"*, OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, https://doi.org/10.1787/cb6d9eca-en. [40]

OECD (2021), *Data-Driven, Information-Enabled Regulatory Delivery*, OECD Publishing, Paris, https://doi.org/10.1787/8f99ec8c-en. [83]

OECD (2020), "Dealing with digital security risk during the Coronavirus (COVID-19) crisis"*, OECD Policy Responses to Coronavirus (COVID-19)*, OECD Publishing, Paris, https://doi.org/10.1787/c9d3fe8e-en. [100]

OECD (2020), *Embracing Innovation in Government - Global Trends 2020: Innovative responces to the COVID-19 crisis*, OECD Publishing, https://trends.oecd-opsi.org/trend-reports/innovative-covid-19-solutions/. [2]

OECD (2020), "The OECD Digital Government Policy Framework: Six dimensions of a Digital Government"*, OECD Public Governance Policy Papers*, No. 02, OECD Publishing, Paris, https://doi.org/10.1787/f64fed2a-en. [3]

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. [65]

OECD (2018), *OECD Regulatory Enforcement and Inspections Toolkit*, OECD Publishing, Paris, https://doi.org/10.1787/9789264303959-en. [82]

OECD (2018), *Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact*, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264305847-en. [84]

OECD (2017), *Recommendation of the Council on Health Data Governance*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433. [103]

OECD (2014), *Recommendation of the Council on Digital Government Strategies*, OECD Publishing, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0406. [1]

OECD (1980), *Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188. [104]

OECD (forthcoming), *Mapping Relevant Data Collection Mechanisms for AI Training*, OECD Publishing. [130]

OECD.AI (2023), *OECD Expert Group on AI Futures – Meeting 2 (18th & 20th September 2023)*, https://wp.oecd.ai/app/uploads/2024/01/Expert-Group-on-AI-Futures-Meeting-2-Summary.pdf. [16]

OECD.AI (2023), *What do you see as the as the most significant potential benefits and risks of AI 10+ years from now?*, https://oecd.ai/en/network-of-experts/ai-futures/discussions/future-benefits-risks. [159]

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. [10]

OECD/UNESCO (2024), *G7 Toolkit for Artificial Intelligence in the Public Sector*, OECD Publishing, Paris, https://doi.org/10.1787/421c1244-en. [12]

Olson, P. (2023), *Don't Go Down That AI Longtermism Rabbit Hole*, https://www.bloomberg.com/opinion/articles/2023-05-19/ai-longtermism-alarmists-are-dragging-us-all-down-existential-rabbit-hole. [118]

Pahlka, J. (2024), *AI meets the cascade of rigidity*, https://www.niskanencenter.org/ai-meets-the-cascade-of-rigidity/. [157]

Passi, S. and M. Vorvoreanu (2022), *Overreliance on AI: Literature review*, https://www.microsoft.com/en-us/research/publication/overreliance-on-ai-literature-review/ (accessed on September 2024). [119]

Peixoto, T., O. Canuto and L. Jordan (2024), *AI and the Future of Government: Unexpected Effects and Critical Challenges*, https://www.policycenter.ma/publications/ai-and-future-government-unexpected-effects-and-critical-challenges. [146]

Pencheva, I., M. Esteve and S. Mikhaylov (2018), "Big Data and AI – A transformational shift for government: So, what next for research?", *Public Policy and Administration*, Vol. 35/1, pp. 24-44, https://doi.org/10.1177/0952076718780537. [42]

Peng, S. et al. (2023), *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot*, https://arxiv.org/abs/2302.06590. [50]

Perry, A. and N. Turner Lee (2019), *AI is coming to schools, and if we're not careful, so will its biases*, https://www.brookings.edu/articles/ai-is-coming-to-schools-and-if-were-not-careful-so-will-its-biases/. [143]

Purdy, M. (2024), *What Is Agentic AI, and How Will It Change Work?*, https://hbr.org/2024/12/what-is-agentic-ai-and-how-will-it-change-work. [19]

Rainie, L. and J. Anderson (2024), *Experts Imagine the Impact of Artificial Intelligence by 2040*, https://imaginingthedigitalfuture.org/wp-content/uploads/2024/02/AI2040-FINAL-White-Paper-2-2.29.24.pdf. [106]

Röttger, P. et al. (2024), "SafetyPrompts: a Systematic Review of Open Datasets for Evaluating and Improving Large Language Model Safety", *arXiv.org*, https://doi.org/10.48550/arXiv.2404.05399. [145]

Russell, S. (2019), *Human Compatible: Artificial Intelligence and the Problem of Control*, Viking. [113]

Saheb, T. (2022), "Ethically contentious aspects of artificial intelligence surveillance: a social science perspective", *AI and Ethics*, Vol. 3/2, pp. 369-379, https://doi.org/10.1007/s43681-022-00196-y. [99]

Salem, A. et al. (2018), *ML-Leaks: Model and Data Independent Membership Inference Attacks and Defenses on Machine Learning Models*, https://arxiv.org/abs/1806.01246. [134]

Santiso, C. (2023), *Public Governance in the Age of Artificial Intelligence*, https://www.chandlerinstitute.org/governancematters/public-governance-in-the-age-of-artificial-intelligence. [39]

Santos, R. et al. (2024), "The use of AI in government and its risks: lessons from the private sector", *Transforming Government: People, Process and Policy*, https://doi.org/10.1108/tg-02-2024-0038. [36]

Sanzogni, L., G. Guzman and P. Busch (2017), "Artificial intelligence and knowledge management: questioning the tacit dimension", *Prometheus*, Vol. 35, pp. 37-56, https://doi.org/10.1080/08109028.2017.1364547. [78]

Sapci, A. and H. Sapci (2019), "Innovative Assisted Living Tools, Remote Monitoring Technologies, Artificial Intelligence-Driven Solutions, and Robotic Systems for Aging Societies: Systematic Review", *JMIR Aging*, Vol. 2/2, p. e15429, https://doi.org/10.2196/15429. [45]

Shane, J. (2019), *You Look Like a Thing and I Love You: How Artificial Intelligence Works and Why It's Making the World a Weirder Place*, https://www.hachettebookgroup.com/titles/janelle-shane/you-look-like-a-thing-and-i-love-you/9780316525220/. [98]

Singh, V. and J. Chobotaru (2022), "Digital Divide: Barriers to Accessing Online Government Services in Canada", *Administrative Sciences*, Vol. 12/3, p. 112, https://doi.org/10.3390/admsci12030112. [140]

Staab, R. et al. (2023), *Beyond Memorization: Violating Privacy Via Inference with Large Language*, https://arxiv.org/abs/2310.07298. [133]

Sunstein, C. (2023), "The use of algorithms in society", *The Review of Austrian Economics*, Vol. 37/4, pp. 399-420, https://doi.org/10.1007/s11138-023-00625-z. [71]

Sun, W., P. Bocchini and B. Davison (2020), "Applications of artificial intelligence for disaster management", *Natural Hazards*, Vol. 103/3, pp. 2631-2689, https://doi.org/10.1007/s11069-020-04124-3. [74]

Tegmark, M. (2017), *Life 3.0: Being Human in the Age of Artificial Intelligence*, Penguin, https://mitpressbookstore.mit.edu/book/9781101970317. [110]

The Economist (2025), *How AI will divide the best from the rest*, https://www.economist.com/finance-and-economics/2025/02/13/how-ai-will-divide-the-best-from-the-rest. [51]

Tse, T. and S. Karimov (2022), *Decision-making risks slow down the use of artificial intelligence in business*, https://blogs.lse.ac.uk/businessreview/2022/05/18/decision-making-risks-slow-down-the-use-of-artificial-intelligence-in-business-1/.

[91]

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.

[9]

UC Berkeley (2021), *Positive AI Economic Futures*, World Economic Forum, https://www.weforum.org/reports/positive-ai-economic-futures.

[125]

UN (2024), *Governing AI for Humanity*, https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

[108]

UN (2022), *E-Government Survey 2022: The Future of Digital Government*, United Nations, https://desapublications.un.org/sites/default/files/publications/2022-09/Web%20version%20E-Government%202022.pdf.

[55]

UNESCO (2019), *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*, https://www.gcedclearinghouse.org/sites/default/files/resources/190175eng.pdf.

[142]

US GAO (2024), *Fraud Risk Management: 2018-2022 Data Show Federal Government Loses an Estimated $233 Billion to $521 Billion Annually to Fraud, Based on Various Risk Environments*, https://www.gao.gov/products/gao-24-105833.

[81]

USGS (2024), *Landsat's Economic Value increased to $25.6 Billion in 2023*, https://www.usgs.gov/news/featured-story/landsats-economic-value-increases-256-billion-2023.

[85]

Valle-Cruz, D. et al. (2020), "Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation", *Government Information Quarterly*, Vol. 37/4, p. 101509, https://doi.org/10.1016/j.giq.2020.101509.

[60]

Valle-Cruz, D., R. Garcia-Contreras and R. Gil-Garcia (2023), "Exploring the negative impacts of artificial intelligence in government: the dark side of intelligent algorithms and cognitive machines", *International Review of Administrative Sciences*, https://doi.org/10.1177/002085232311870.

[96]

Williams, C. (2025), *There will be no immediate productivity boost from AI*, https://www.economist.com/the-world-ahead/2024/11/20/there-will-be-no-immediate-productivity-boost-from-ai.

[29]

Wirjo, A. et al. (2022), *Artificial Intelligence in Economic Policymaking*, Policy Brief No. 52, https://www.apec.org/docs/default-source/publications/2022/11/artificial-intelligence-in-economic-policymaking/222_psu_artificial-intelligence-in-economic-policymaking.pdf.

[63]

Xu, G., M. Xue and J. Zhao (2023), "The Relationship of Artificial Intelligence Opportunity Perception and Employee Workplace Well-Being: A Moderated Mediation Model", *International Journal of Environmental Research and Public Health*, Vol. 20/3, p. 1974, https://doi.org/10.3390/ijerph20031974.

[44]

Zehnle, M., C. Hildebrand and A. Valenzuela (2025), "Not all AI is created equal: A meta-analysis revealing drivers of AI resistance across markets, methods, and time", *International Journal of Research in Marketing*, https://doi.org/10.1016/j.ijresmar.2025.02.005. [156]

Zhang, Z., L. Wang and C. Lee (2023), "Recent Advances in Artificial Intelligence Sensors", *Advanced Sensor Research*, Vol. 2/8, p. 2200072, https://doi.org/10.1002/adsr.202200072. [77]

Notes

[1] https://www.oecd.org/en/topics/policy-issues/digital-government.

[2] The United States alone has catalogued more than 2 000 use cases in civilian federal government agencies (https://github.com/ombegov/2024-Federal-AI-Use-Case-Inventory). Similarly, the European Commission has catalogue more than 1 300 (https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/cases). More than 700 use cases have been catalogued in Latin American and Caribbean (LAC) governments (https://sistemaspublicos.tech/sistemas-de-ia-en-america-latina).

[3] See https://trends.oecd-opsi.org, https://cross-border.oecd-opsi.org, (OECD, 2023[38]), and (OECD, 2024[105]).

[4] See https://oecd-opsi.org/innovation-tag/artificial-intelligence-ai and https://oecd.ai/en/dashboards/policy-initiatives?orderBy=startYearDesc&page=1&terms=&initiativeTypeIds=123.

[5] These findings are based on a survey of 1 000 senior executives from 20 sectors across 59 countries in Asia, Europe and North America.

[6] Unless otherwise cited, the sections below on the benefits of AI in the public sector are based on analysis of the functions of government and use cases presented in Chapter 5 of this report.

[7] "Noise" is also a term used in ML to mean "random or unpredictable fluctuations in data that disrupt the ability to identify target patterns or relationships. The result is decreased accuracy or reliability of a model's predictions or output" (DataHeroes, 2023[160]). This is not the concept of noise discussed in this chapter, which focuses on factors that can influence humans.

[8] For more information on this topic, see https://www.oecd.org/en/topics/behavioural-science and https://oecd-opsi.org/work-areas/behavioural-insights.

[9] See also https://www.oecd.org/en/topics/sub-issues/privacy-enhancing-technologies.

[10] See the MIT AI Risk Repository, a living database of over 1 000 AI risks (https://airisk.mit.edu). The OECD AI Incidents Monitor (AIM) documents AI incidents and hazards to help policymakers, AI practitioners, and all stakeholders worldwide gain valuable insights into the risks and harms of AI systems (https://oecd.ai/incidents).

[11] See https://oecd.ai/dashboards and (OECD, 2024[14]).

[12] Overfitting refers to cases where the algorithm is too specific to the extent that it captures and focuses too much on noise and anomalies (Berryhill et al., 2019[8]). During the training phase, an overfitting model may achieve a high level of accuracy and problems may go unnoticed. However, once the trained model is exposed to new data, accuracy can drop severely.

[13] The index does not distinguish between legitimate and illegitimate uses of AI surveillance techniques. Rather, the purpose of the research is to show how new surveillance capabilities are transforming governments' ability to monitor and track individuals or groups.

[14] Digital Taylorism refers to the modern adaptation of Frederick Winslow Taylor's principles of scientific management, utilising digital technologies to monitor and control employee activities with the goal of enhancing efficiency and productivity. This approach involves breaking down complex tasks into simpler components, standardising workflows, and employing data-driven methods to oversee and evaluate worker performance. While it aims to optimise organisational operations, critics argue that it may lead to decreased worker autonomy and increased surveillance in the workplace (Grzegorzek, 2024[161]).

[15] https://oecd.ai/site/data-privacy.

[16] The OECD Directorate for Science, Technology and Innovation (STI) has a dedicated line of work on digital security. See https://www.oecd.org/en/topics/policy-issues/digital-security for more information.

[17] An **AI incident** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems directly or indirectly leads to any of the following harms: (a) injury or harm to the health of a person or groups of people; (b) disruption of the management and operation of critical infrastructure; (c) violations of human rights or a breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights; or (d) harm to property, communities or the environment. An **AI hazard** is an event, circumstance or series of events where the development, use or malfunction of one or more AI systems could plausibly lead to an AI incident, i.e. any of the following harms: (a) injury or harm to the health of a person or groups of people; (b) disruption of the management and operation of critical infrastructure; (c) violations to human rights or a breach of obligations under the applicable law intended to protect fundamental, labour and intellectual property rights; or (d) harm to property, communities or the environment. For more information, see https://oecd.ai/incidents-methodology and (OECD, 2025[162]; 2024[163]).

2 Trends and early lessons from the use of AI across functions of government

This chapter synthesises 200 AI use cases across 11 government functions. It finds using AI is a priority for governments, but adoption is fragmented and uneven. Use concentrates on public facing services and internal operations, with fewer examples in policymaking. Governments pursue productivity, responsiveness and accountability, yet efforts to empower external actors are limited. Maturity varies by function and technology, with long-standing rules-based systems, selective machine learning and limited generative AI. Every use case can pose operational, ethical, resistance or exclusion risks if not trustworthy, underscoring the need for strong data foundations and coherent governance.

Key messages

- The OECD analysed 200 use cases across 11 government functions. It found that while AI is a priority for most governments, efforts are not systematic.
- This analysis helped to better understand the current state of AI in government and to identify overarching trends. These trends are that:
  o AI is unevenly distributed across government functions
  o AI is most used for public-facing public service activities and internal operations
  o Governments are using AI to pursue a variety of potential benefits
  o Some government functions are more mature regarding governing and adopting AI
  o Different functions of government have different contexts and needs.
- The OECD found that every one of the 200 use cases analysed could post one or more types of risk (operational, ethical, public resistance or exclusion) if not designed and used in a trustworthy way.
  o These risks vary across use cases and functions of government; therefore, it is important to acknowledge the main drivers of risks in each use and field.
  o While governments are vigilant of several AI risks, some receive less focus.

OECD analysis of 200 use cases across 11 government functions

In its latest cross-cutting work on AI in government, the OECD (2024[1]) found the need for the systematic collection, documentation and analysis of AI use cases to monitor trends on policy options across countries. The OECD also found that more and better evidence of the impact of AI on governments will help ensure the technology is used for optimal impact. Ease of access to such evidence as well as information on policies, practice and use of AI in government could promote progress in trustworthy AI adoption, structured dialogue and exchanges among countries. Overall, there is a need for a holistic, systems approach to maximise the value of AI in government, including establishing enablers, guardrails and engagement mechanisms.

To help address these needs, this chapter considers and builds upon OECD and other relevant research to better understand the current state of AI in government and to illuminate overarching trends. This chapter analyses and synthesises 200 AI use cases spanning 11 government functions, as listed in Table 2.1 and discussed in-depth in Chapter 5.[1] These use cases were identified through, and the findings discussed in this chapter informed by, desktop research, OECD meetings and discussions with public officials in relevant OECD working parties and networks, and ongoing data collections from the OECD Observatory of Public Sector Innovation (OPSI) and the OECD.AI Policy Observatory.[2]

Based on this methodology, the findings in this chapter are not generalisable to the broader universe of AI efforts in government. In addition, adoption of AI in government will vary across countries, depending on their national realities and levels of AI readiness. The findings do provide, however, observations rooted in real-world practice, the latest research and policymakers' present points of view. In doing so, the chapter seeks to take the pulse of current activities and their characteristics, as well as potential gaps where there may be untapped potential or need for further research.

In the coming months, the OECD will establish a living global repository of relevant initiatives and case studies as part of the OECD.AI Policy Observatory.

Table 2.1. Functions of government analysed for *Governing with AI*

Category	Function	Scope of analysis
Government policy functions	Tax administration	OECD experts in each function of government leveraged OECD and external research and analysed 200 use cases to determine: - current state of play, including government actions and benefits pursued, - relevant challenges and how governments are managing them, and - untapped potential and the way forward.
	Public financial management	
	Regulatory design and delivery	
Key government processes	Civil service reform	
	Public procurement	
	Fighting corruption and promoting public integrity	
	Policy Evaluation	
	Civic participation and open government	
Government services and justice functions	Public service design and delivery	
	Law enforcement and disaster risk management	
	Justice administration and access to justice	

AI is a priority, but efforts are not systemic

In all, 48 countries and the European Union (EU) have adhered to the OECD AI Principles (Table 2.2), committing to promoting the trustworthy design, development, deployment and use of AI, including in the public sector. The OECD is tracking and reporting on the implementation of these principles over time (2023[2]; 2021[3]). Findings suggest that government are less focused on using AI than on their efforts to promote trustworthy AI adoption in the broader economy and society.

Table 2.2. OECD AI Principles

	Principle	Description
Value-based principles	**Inclusive growth**, **sustainable development** and **well-being** (Principle 1.1)	Highlights the potential for trustworthy AI to contribute to overall growth and prosperity for all – individuals, society, and planet – and advance global development objectives.
	Respect for the **rule of law**, **human rights** and **democratic values**, including fairness and privacy (Principle 1.2)	AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and should include appropriate safeguards to ensure a fair and just society.
	Transparency and **explainability** (Principle 1.3)	About transparency and responsible disclosure around AI systems to ensure that people understand when they are engaging with them and can challenge outcomes.
	Robustness, **security** and **safety** (Principle 1.4)	AI systems should function in a robust, secure and safe way throughout their lifetimes, and potential risks should be continually assessed and managed.
	Accountability (Principle 1.5)	Organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the OECD's values-based principles for AI.
Recommendations for policymakers	**Investing** in AI research and development (Principle 2.1)	Governments should facilitate public and private investment in research & development to spur innovation in trustworthy AI.
	Fostering an inclusive **AI-enabling ecosystem** (Principle 2.2)	Governments should foster accessible AI ecosystems with digital infrastructure and technologies, and mechanisms to share data and knowledge as well as ensure the quality of such information.
	Shaping and enabling **interoperable governance and policy environment** for AI (Principle 2.3)	Governments should create a policy environment that will open the way to deployment of trustworthy AI systems.
	Building **human capacity** and preparing for labour market transformation (Principle 2.4)	Governments should equip people with the skills for AI and support workers to ensure a fair transition.
	International co-operation for trustworthy AI (Principle 2.5)	Governments should co-operate across borders and sectors to share information, develop standards and work towards responsible stewardship of AI.

Source: https://oecd.ai/en/ai-principles.

Governments are realising the potential of AI in public administrations and making it a strategic priority. **Almost all OECD countries have put in place strategies and agendas for AI that establish a high-level vision and approach for its use in government.** These are mainly embedded in broader national AI strategies. However, countries like Canada (2025[4]), Switzerland (2025[5]) and Uruguay (2021[6]) have developed dedicated strategies. Many governments have also sought to convert strategy into practice through cross-cutting or domain-specific policies and initiatives. Such efforts are discussed further in Chapter 4. In addition, many governments have adopted AI through hands-on use and, increasingly, custom development.

Progress has been made since the OECD began exploring AI in government in 2019; however, **efforts to date are limited and not systematic.** The potential opportunities of AI in government are significant but not easy to attain. Governments in two-thirds of OECD countries have started to explore the use of AI for internal efficiency by enhancing processes. Yet more progress is needed not only in using AI for other purposes, such as improving policies, but also in building the foundational components needed for AI in government to flourish (OECD, 2024[1]). In addition, the review of use cases suggests a proliferation of AI tools, with implementation often occurring in a piecemeal manner. These efforts are frequently undertaken without overarching governance mechanisms to steer initiatives across sectors or government as a whole, or to draw lessons from their implementation. As a result, the potential for coordinated learning, scaling and impact remains limited. Establishing robust governance frameworks could help ensure AI systems are deployed in a cohesive, efficient and accountable way, aligned with strategic priorities and public values.

The sections below seek to uncover the current status of AI use in government, including key patterns and trends among early adopters. In doing so, these sections seek to identify the extent to which governments extend beyond principles to take action in using AI, what results and outcomes they are achieving and what limitations they face.

General trends in government AI use cases

Uneven distribution of AI use across government functions

The 2023 OECD Digital Government Index (DGI) ([7]) found that while some countries have deployed a wide range of initiatives to enhance their capacity to use AI in government, implementation is still a challenge across most countries. In digging deeper into the use cases analysed for this report, OECD analysis suggests government AI efforts may cluster around public service design and delivery, civic participation and open government, and justice administration and access to justice. Conversely, only few initiatives in functions such as policy evaluation were identified (Figure 2.1).

Figure 2.1. Use cases are most present in public service, civic participation and justice functions
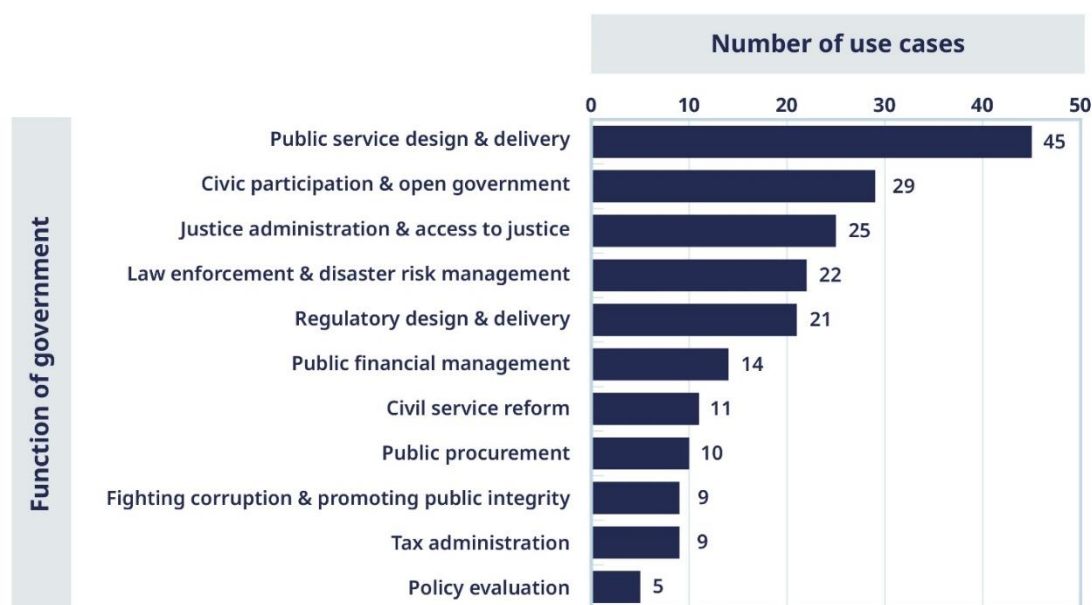
Source: OECD analysis of identified use cases.

There are several potential explanations for this distribution:

- Public service design and delivery extends horizontally across many different types of organisations and issue areas, making it more prevalent in terms of total use cases than more vertical functions of government, such as tax administration.

- Civic participation and open government's prevalence may be partially because it is unencumbered by many of the risks (Chapter 1) and challenges (Chapter 3) faced by other functions. For instance, concerns around data access and security are largely non-applicable because the point of such engagement is generally to gather data on issues and questions that are public by nature. In addition, government teams engaging in civic participation are often among the more innovative groups in the public service, and thus, perhaps more prone to embracing new technological approaches.

- The policy functions most represented tend to be public facing, potentially suggesting a focus on areas with immediate visibility to citizens. Factors comprising this could be more demands from citizens, as well as a desire among government and political leaders to visibly demonstrate value.

- Some functions face barriers or complexities, such as stringent rules on data access and sharing in tax administration and requirements for thorough audit trails in public integrity.

- Some functions appear to be more mature than others pertaining to AI readiness, including their underlying foundations for AI, such as sufficient and quality data, as discussed below.

- Some functions may have pre-existing structures and processes that cannot be easily substituted or complemented by AI systems.

The prevalence of AI use in justice and the related function of law enforcement is particularly interesting. In general, the OECD (2024[1]) has encouraged governments to aim for low-hanging fruit in their initiation of AI — focusing on areas that represent high-benefit, low-risk uses of AI. The use of AI in justice and law enforcement can be high benefit but also high risk. One reason for the prevalence of use cases in some of these areas may be the prevalence of more comprehensive and structured data. Another reason may align with the sheer volume of tasks required for some of these functions. In the case of justice administration and access to justice, justice systems worldwide often operate under tight resource

constraints, including limited budget and court staff, even as the volume of cases continues to grow (Harvard Kennedy School, 2023[8]; Columbia University, 2020[9]). This mismatch has led to chronic case backlogs in many jurisdictions, creating intense pressure on court administrators to explore technologies that can boost productivity and mitigate the backlog problem. This might be reflected in the higher number of AI use cases in justice linked to internal operations, compared to other government functions (Figure 2.3).

An additional reason for this distribution may be the data available for review, with the OECD tending to identify, or governments tending to submit, information on initiatives in some functions more than others. This seems somewhat tempered by validation in OECD discussions and reviews by the OECD Working Party on Senior Digital Government Officials (E-Leaders), as well comparisons with results from larger databases. Regarding the latter, the data collection for this report aligns with the trends seen in the EU and Latin America and the Caribbean (LAC), as recorded by the European Commission (EC) Public Sector Tech Watch observatory (2025[10]) and the "AI Systems in the Public Sector in LAC" database (Muñoz-Cadena et al., 2025[11]) (Figure 2.2). In the EU, the top three functions identified in their repository of nearly 1 500 AI uses, as of 31 March 2025, are general public services, economic affairs, and public order and safety.[3] The public order and safety category includes use cases in both law enforcement and justice administration. The economic affairs category is mainly represented by use cases in sectors like transport, agriculture and energy. About 70% of economic affairs AI use cases are related to regulation and targeted public services and engagement. LAC exhibits a comparable trend in the top three functions within the reviewed repository of approximately 700 AI systems, as updated on 19 March 2025.

Figure 2.2. EU and LAC follow a similar trend with the AI use cases sample collected for this report

Percentage of use cases categorised according to the United Nation's Classification of the Functions of Government
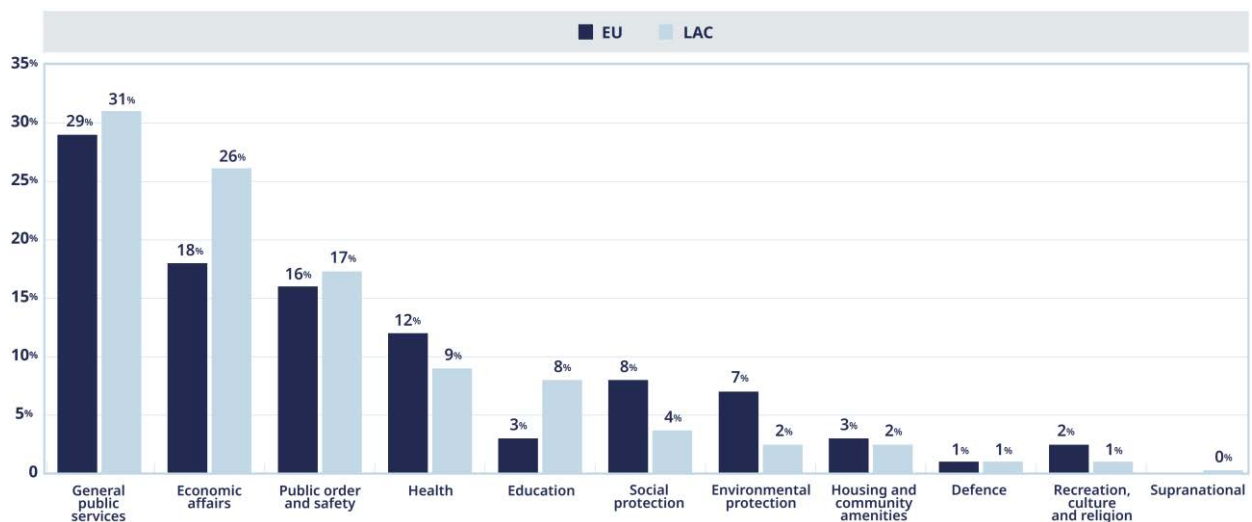
Source: Data analysed by the OECD from (EC, 2025[10]; Muñoz-Cadena et al., 2025[11]).

A final factor that may also influence the results is a variation on the "AI effect", whereby "as soon as [AI] researchers achieve a milestone long thought to signify the achievement of true artificial intelligence, e.g. beating a human at chess, it suddenly gets downgraded to not true AI" (Bailey, 2016[12]). In discussions with the E-Leaders working party, delegates have suggested that narrow and traditional applications of AI may have become so integrated or commonplace, they no longer trigger external reporting or a response to data collection efforts. This could potentially occur more in areas with longstanding use of such systems, such as tax administration, resulting in less representation in examined initiatives. It is difficult to determine the extent to which this may occur; however, the analysis for this report did identify and include many such use cases.

AI is most used for public-facing public service activities and internal operations

Across the 11 core functions covered by this report, governments are using AI in four general activities: public-facing service delivery, internal operations, internal and external oversight activities and assisting policymaking. The use of AI is most common in public-facing service delivery. Activities related to internal operations are not far behind (Figure 2.3). Internal and external oversight activities and assisting policymaking were not as prevalent. This is not unexpected, as service delivery and internal operations constitute the majority of what government organisations do. Oversight activities are important, though they are more often limited to certain offices or teams. AI use in policymaking activities is also not as prevalent throughout governments. This finding aligns with previous measurements by the OECD DGI (2024[7]), which also indicated that governments could make more effort in this area. Many remain cautious or lack the necessary skills to incorporate AI into decision-making processes.

AI's prevalence in government activities varies according to the nature of each core function. In the function of public service design and delivery, most AI applications naturally involve public-facing government-citizen interactions, with some use cases also addressing internal operations regarding how public services are designed or delivered. The significant number of service-related use cases in the justice sector, comprising almost three quarters (16 of 25 cases, noting that one case may address more than one activity) of the documented instances in that function, indicates that this function has prioritised responsiveness to citizens, along with enhancing efficiency in its internal operations. This focus may be influenced by greater societal demands, pressure to reduce case backlogs and manage scarce resources. Functions such as civic participation and regulatory design and delivery encompass most of the use cases related to policymaking activities. Related use cases include supporting the processing of evidence and stakeholder inputs for policy formulation, and various applications aiding decision making through analytics, simulations or forecasting.

Figure 2.3. AI use cases per core function and government activity
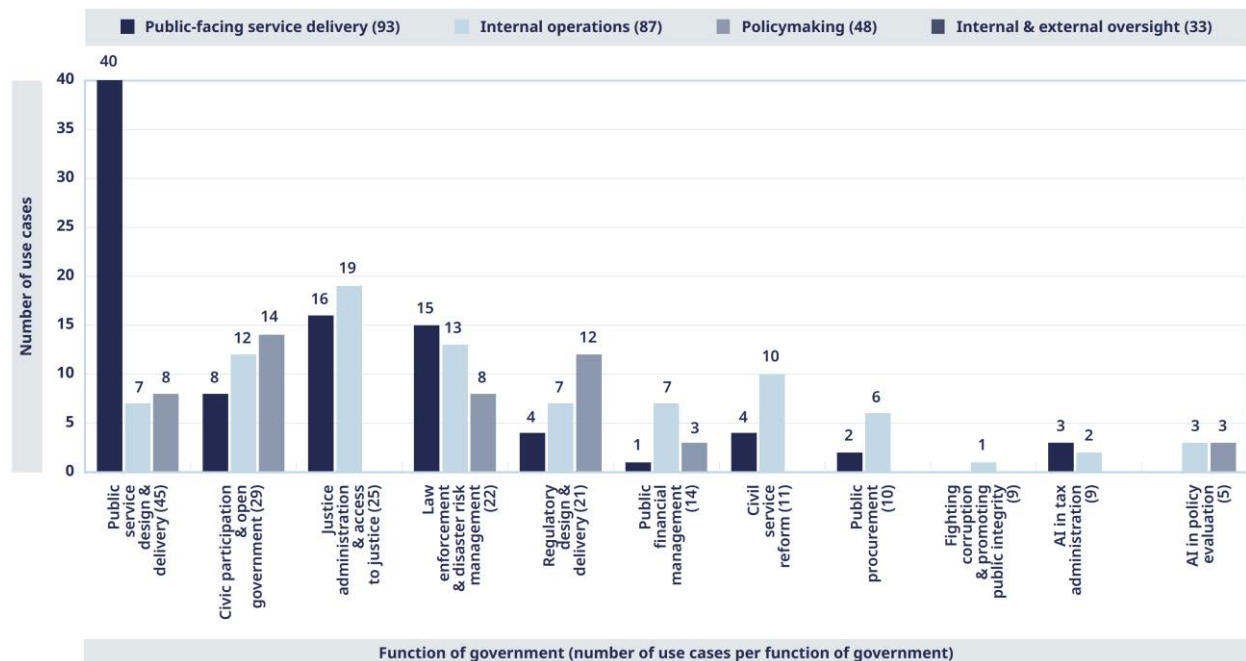
Note: The four activities in this figure are not mutually exclusive (e.g. one AI use case could seek to both improve internal operations and service delivery). Thus, the sum of activities is greater than the total number of use cases.
Source: OECD analysis of identified use cases.

These results are generally in line with findings from the EU (2025[10]) and LAC (2025[11]) databases. Although there is not a uniform methodology for categorising government activities, these and OECD's databases do tend to show a common trend, particularly around services. Both the EU and LAC databases include a classification of use cases under government activities (Figure 2.4), where "public services and engagement" represent a significant share of all use cases, being the most prominent in EU and the second most in LAC. However, in LAC, most use cases are categorised as part of "enforcement" activities, which include predictive enforcement, registration and data notarisation, smart recognition, supporting inspections, among others. These processes would generally coincide with OECD's internal operations and oversight (internal and external) categories. This means that the "internal management" category in EU and LAC, which is generally less representative for these databases, is not the only one containing uses cases that the OECD could categorise as "internal operations". Therefore, it is not possible to conclude whether the OECD dataset contains a higher share of use cases belonging internal operations of government, compared to EU and LAC. Finally, it is worth highlighting that processes under the "analysis, monitoring, and regulatory research" category in EU and LAC generally match with OECD's "policymaking" category and have a similar share (about 20-30% of all use cases).

Figure 2.4. Public services and engagement represent an important share of use cases under government processes in the EU and LAC

Percentage of use cases categorised by how the technology supports government decision-making and implementation
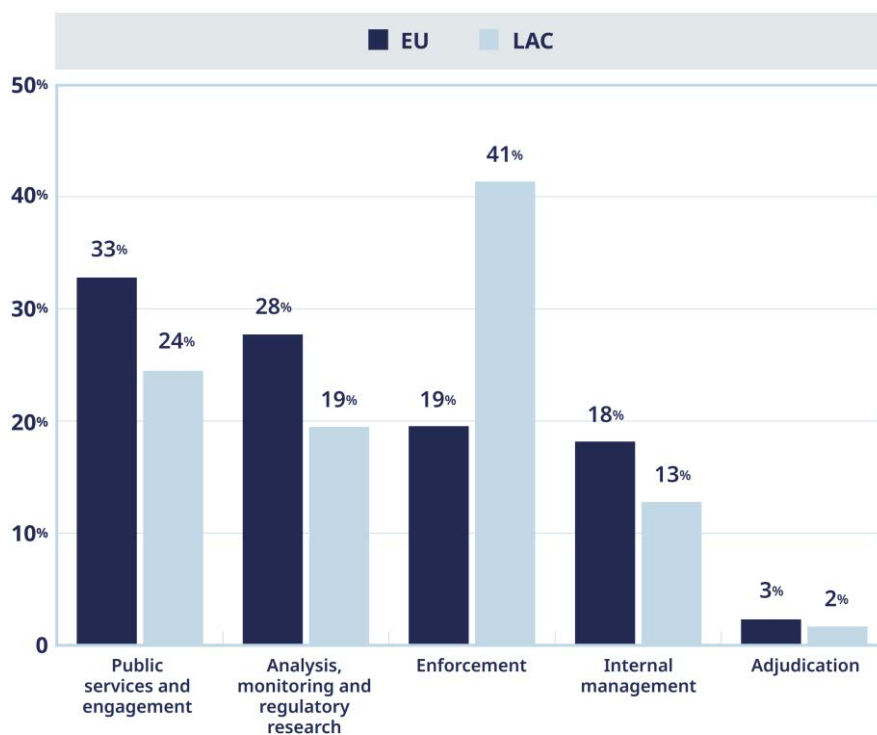
Source: Data analysed by the OECD from (EC, 2025[10]; Muñoz-Cadena et al., 2025[11]).

Governments are using AI in pursuit of a variety of potential benefits

The use cases analysed by the OECD have the potential to address all the AI benefits introduced in Chapter 1, and most use cases have the potential to yield multiple benefits. However, certain benefits receive stronger focus from governments than others (Figure 2.5):

- About six out of every 10 of the examined use cases seek to contribute to the automation, streamlining and tailoring and personalisation of processes and services, particularly within justice, public services, civic participation and regulatory design and delivery.
- Nearly half of all use cases seek to enhance decision-making, sense-making and forecasting, with most concentrated in public services, regulation and civic participation.
- About a third of the use cases have the potential to improve accountability and anomaly detection, mainly within law enforcement and disaster risk management, civic participation, fighting corruption and promoting public integrity, public procurement and regulation.
- A small proportion of use cases have the potential to unlock opportunities for external stakeholders, such as citizens, civil society organisations and businesses, especially in civic participation, access to justice and disaster risk management.

The lack of emphasis on unlocking opportunities for external stakeholders through AI as a good for all stands out as a potential gap. AI experts have suggested that this type of empowerment is important and that governments could take more action to seize it (OECD, 2024[13]). However, such efforts could potentially be more prevalent in areas not covered by this report (e.g. specific sectors, such as agriculture or education). While the results of this benefit may be less directly felt by governments, they can pay dividends through strengthened trust in government or even economic gains.

Figure 2.5. Potential benefits of AI use cases across functions of government

Percentage of use cases for the corresponding function of government
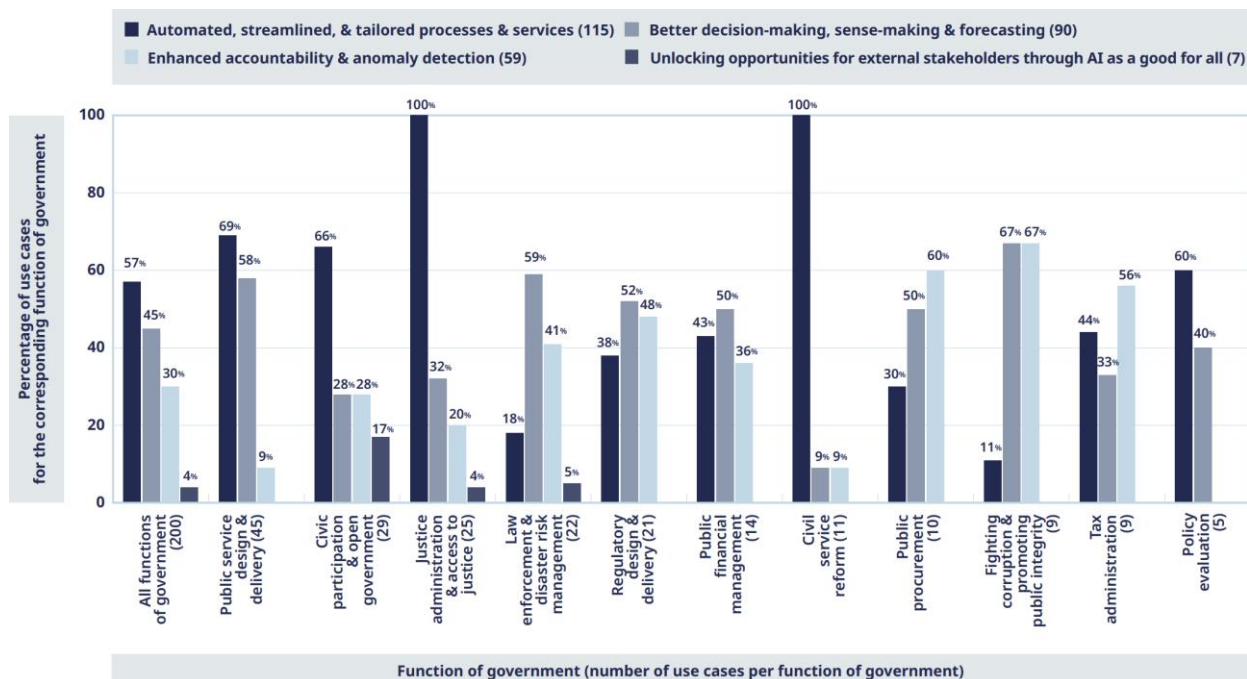
Note: The potential benefits in this figure are not mutually exclusive (i.e. one use case may have the potential to yield more than one type of benefit). Thus, the sum of potential benefits observed is greater than the total number of use cases.
Source: OECD analysis of identified use cases.

When examining the specific benefits within the four general activity categories mentioned above, more detailed insights into the direct gains governments aim to achieve through the use of AI can be ascertained

(as shown in Figure 2.6). The sections below further detail these benefits and provide examples of some of the use cases that informed such trends.

Figure 2.6. Specific benefits of AI use cases
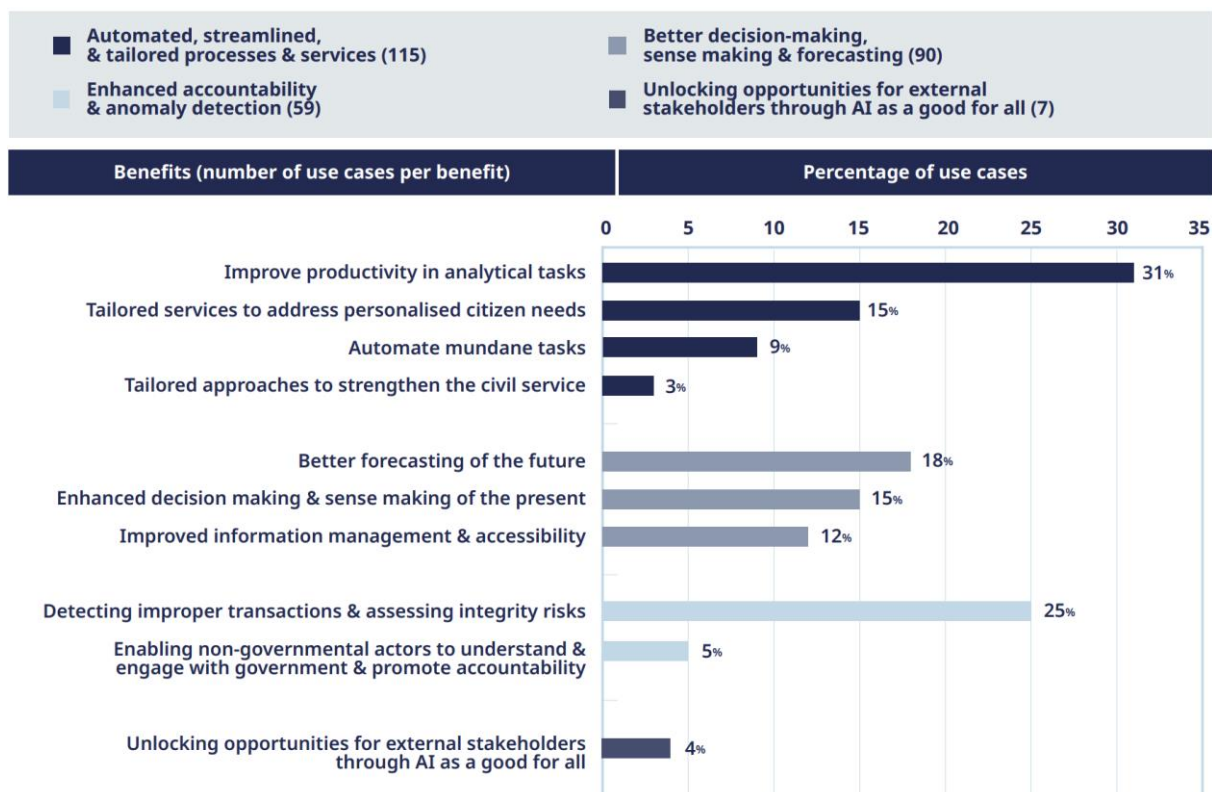
Note: The potential benefits in this figure are not mutually exclusive (i.e. one use case may have the potential to yield more than one type of benefit). Thus, the sum of potential benefits observed is greater than the total number of use cases.
Source: OECD analysis of identified use cases.

Automated, streamlined and tailored processes and services

About a third (31%) of the analysed use cases aim to improve productivity in analytical tasks. To a lesser extent, 15% of use cases represent government efforts to use AI to tailor services to address personalised citizen needs. This relatively lower adoption of AI for personalisation could be partly attributed to data governance limitations or restrictions due to the large volume of personal data required for such applications (see Chapter 3 for discussion on implementation challenges). This appears to be a gap warranting further analysis. Interestingly, the automation of mundane tasks comes in at 9% of the analysed use cases. This is contrary to conventional expectations of AI primarily being used for automating repetitive tasks that require little analytical consideration (Figure 2.6). While a case review methodology is not fully generalisable to the universe of AI in government, this suggests a potential shift of AI's use towards enhancing more complex decision-making processes and supporting more specialised work of public servants and policymakers. Yet, it could also suggest governments are not fully capitalising on AI with respect to repetitive tasks, which public servants spend a significant amount of time on, and for which tremendous efficiencies can be made through AI (The Alan Turing Institute, 2024[14]; Berryhill et al., 2019[15]).

Table 2.3 provides examples of how AI is being used for these purposes. Use cases intended to improve productivity in analytical tasks include uses like estimating compliance costs in regulatory impact

assessments (Germany), analysing and scoring candidates' recorded responses in certain recruitment processes (United Kingdom), or supporting government staff with common procurement queries (North Carolina, United States). Automation in repetitive tasks that require less intellectual engagement encompasses various domains, including repetitive judicial tasks or financial and HR processes. This is the case of Prometea in Argentina, the AI Litigation Project in Brazil or Finland's use of RPA and AI in financial management. Uses aimed at tailoring services and personalisation can be seen in functions such as public services, tax, regulation or justice. For example, the Public Employment Service in Sweden uses BÄR to tailor job-finding support, optimising resource allocation through personalised training and guidance recommendations. In the case of Singapore, the tax authority developed a chatbot to enhance self-service by assisting taxpayers with inquiries and payments. Finally, AI is being used to strengthen civil service hiring and professional development programmes, such as the Australian Public Service Commission trial to use AI to expedite the design, structuring and deployment of digital skills training; or the Spanish National Institute for Public Administration use to transform how civil servants access and use learning resources by improving searchability and recommendations of relevant materials.

Table 2.3. Examples of AI for automated, streamlined and tailored processes and services

Country	Initiative	Description	Sub-benefit	Function
Argentina	Prometea and ChatGPT in the justice sector	The Public Prosecution Service of Buenos Aires adopted Prometea in 2017 to automate repetitive judicial tasks and expedite case proceedings. In 2024, it began to also is exploring explore the use of ChatGPT to analyse legal cases and draft decisions. This AI tool reduces sentencing drafting time from an hour to 10 minutes, increasing efficiency in case management.	Automate mundane (and recently, analytical) tasks	Justice (Box 5.62)
Australia	AI to generate an online learning	The Australian Public Service Commission (APSC) trialled accelerating course creation for public servants by using AI to design, structure and deploy digital skills training in minutes rather than weeks. The pilot trialled feeding in controlled materials to generate course outlines and quizzes and refine content through feedback loops.	Tailored approaches to strengthen the civil service	Civil service reform (Box 5.22)
Brazil	AI Litigation Project	Brazil's tax courts use AI to group similar tax appeal cases, assigning them to the same officers for faster processing. Initial trials demonstrated high accuracy, significantly reducing case backlog and improving decision speed.	Automate mundane tasks	Tax administration (Box 5.2)
Finland	RPA and AI in Financial Management	Finland leverages a tool that automates financial and HR processes through RPA and AI, optimising tasks such as invoice processing. Its structured automation strategy improves scalability and efficiency.	Automate mundane tasks	Public finance
Germany	AI for regulatory impact assessments	Germany's Federal Statistical Office is exploring the use of AI to estimate compliance costs in regulatory impact assessments. AI identifies relevant legal text passages and predicts cost implications, allowing officials to focus resources on complex cases.	Improve productivity in analytical tasks	Regulation (Box 5.13)
Singapore	Chatbot for taxpayer services	Singapore's tax authority developed a chatbot using AI and NLP to assist taxpayers with inquiries and payments. The system enhances self-service options, reducing administrative workload and improving user satisfaction.	Tailored services to address personalised needs	Tax administration (Box 5.4)
Spain	Knowledge graph	The National Institute for Public Administration (INAP) AI-enhanced knowledge graph transforms how civil servants access and use vast learning resources. By creating a "resource bank" that improves searchability and recommends relevant materials, INAP enables public officials to efficiently find and apply critical knowledge.	Tailored approaches to strengthen the civil service	Civil service reform
Sweden	BÄR	The Public Employment Service uses BÄR, an AI tool within the Prepare and Match program, to tailor job-finding support. By analysing jobseekers' profiles and predicting employment chances, it guides decisions and optimises resource allocation through personalised training and guidance recommendations.	Tailored services to address personalised needs	Civil service reform (5.43)
United Kingdom	Outmatch	The UK tax authority (HMRC) employs Outmatch to automate junior role recruitment by analysing and scoring candidates' recorded responses. This speeds up high-volume hiring while ensuring consistency in	Improve productivity in analytical tasks	Civil service reform (Box 5.20)

		evaluation.		
United States	Chatbot to support procurement	North Carolina's IT department introduced a 24/7 AI-powered chatbot to support government staff with common procurement queries. It provides instant answers, streamlining processes and reducing wait times.	Improve productivity in analytical tasks	Public procurement

Better decision-making, sense-making and forecasting

The use of AI for enhanced decision-making and sense-making of the present was measured across 18% of cases, while 15% of use cases aimed at better forecasting of the future, and 12% at improving information management and accessibility to support these activities (Figure 2.6). Such uses of AI not only support policymaking processes — which are indeed a minority when it comes to government AI efforts (OECD, 2024[7]) — but also contribute to smarter policy implementation and internal operation, and better quality and pertinence of service design and delivery.

Table 2.4 provides some examples of how AI is being used for these purposes. Governments are using open-source tools like Polis to make better sense of the present, specifically in deliberative exercises, where it clusters public opinions and identifies areas of consensus. Other uses allow governments to better optimise decision-making. An example is Korea's dBrain+, which analyses real-time financial management data and integrates risk assessment, budget management and performance evaluation tools. Most forecasting use cases aim to predict certain conditions to take decisions in advance and pre-position resources — such as predicting slippery conditions for winter road maintenance in Belgium or forecasting the likelihood of wildfires in Canada — to improve risk mitigation and authorities' response times. Some other forecasting uses aim to simulate alternative scenarios. One example is Helsinki's (Finland) use of UrbanistAI to generate visualisations of alternative urban planning scenarios in order to support consensus-building among stakeholders. Finally, the use of AI to improve information management and accessibility can take the form of tools available for stakeholders both inside and outside of government to access vast amounts of data. Examples include the European Parliament's search tool that allows users to analyse over 20 years of parliamentary documents, or the Netherland's Court of Audit GenAI pilot platform for analysing public audit reports. It might also take the form of support tools to quickly retrieve accurate government information and ensure reliable responses in customer support services, such as the virtual assistants *Caddy* in the UK, and *Albert* in France. While those assistants span domains, some are focused on specific areas, such as France's *Sofia* conversational agent for ecological information.

Table 2.4. Examples of AI for better decision-making, sense-making and forecasting

Country	Initiative	Description	Sub-benefit	Function
Belgium	AI for road safety	Belgium predicts slippery conditions and optimises resource allocation for winter road maintenance. By analysing weather and traffic data, the tool helps authorities proactively deploy de-icing measures, improving road safety and reducing accidents.	Better forecasting of the future	Public services; related to disaster management
Canada	Anticipating wildfire risks	Alberta's AI wildfire prediction system forecasts the likelihood of wildfires across the province's protected forests using historical fire, weather and ecological data. The tool assists authorities in pre-positioning resources, improving response times and mitigating risks.	Better forecasting of the future	Law enforcement and disaster management (Box 5.54)
European Union	AI for examining parliamentary documents	The European Parliament's search tool enables citizens and policymakers to efficiently analyse over 20 years of parliamentary documents, including 38 000 motions for resolutions and parliamentary questions. By automating information retrieval, the AI system improves accessibility and facilitates informed decision-making.	Improved information management/ accessibility	Civic participation and open government (Box 5.35)
Finland	UrbanistAI	The City of Helsinki used UrbanistAI to generate visualisations of alternative urban planning scenarios, helping citizens and local businesses engage in discussions about pedestrianising key streets. With AI-generated renderings, the tool supported consensus-building among stakeholders.	Better forecasting of the future	Civic participation (Box 5.38)
France	Albert and	Albert is a GenAI tool developed to assist public administration employees	Improved	Public services

	Sofia	in responding to citizen inquiries. The tool helps civil servants search for regulations, summarise information and draft responses, while human agents verify the final output. Sofia is a conversational agent that facilitates access to the Ministry of Ecological Transition's scientific and technical knowledge.	information management accessibility	(Box 5.46)
International	Polis	Polis is an AI-powered open-source platform designed to facilitate large-scale deliberative processes by clustering public opinions and identifying consensus statements. It has been used in multiple countries to inform climate policy, referendum debates, municipal decision-making and political party platforms.	Enhanced decision-making and sense-making of the present	Civic participation (Box 5.36)
Korea	dBrain+	dBrain+ is an AI-driven financial management system that analyses real-time economic, fiscal and financial data to optimise public finance decision-making. It integrates risk assessment, budget management and performance evaluation.	Enhanced decision-making and sense-making of the present	Public finance (Box 5.8)
Netherlands	GenAI platform on public audit work	The Netherland's Court of Audit, a public GenAI platform is currently being piloted to allow citizens and other stakeholders to roam through public reports and find answers and sources to their questions on public audit work.	Improved information management/ accessibility	Fighting corruption and promoting integrity
United Kingdom	Caddy	Caddy, an AI-powered assistant developed in the UK, supports customer service agents by quickly retrieving accurate government information. With a human-in-the-loop validation system, Caddy ensures reliable responses while improving efficiency in handling citizen inquiries.	Improved information management/ accessibility	Public services

Enhanced accountability and anomaly detection

Regarding uses for enhanced accountability and anomaly detection, 25% of the analysed use cases focused on detecting improper transactions and assessing integrity risks, and 5% on improving governments' ability to engage non-governmental actors and promote accountability (Figure 2.6). The former covers use cases related to oversight, preventive controls, and risk assessment and management. Those are generally linked to the core mandates of some specific functions of government, such as those responsible for fighting corruption and promoting public integrity or enforcing regulatory compliance. The use cases that better connect government with non-governmental actors, ultimately contributing to greater accountability and responsiveness, are often related to the civic participation and transparency practices used in various government functions and organisations.

Table 2.5 provides some examples of how AI is being used to pursue this benefit. Some uses focus on prioritising cautionary actions based on the analysis of patterns and statistical anomalies. For example, Portugal's Court of Audit uses AI to detect critical and priority cases in public procurement that might require concentrating audit efforts. In Chile, AI is used in the country's public procurement platform to detect irregularities and improve compliance monitoring. Other uses are intended to detect loopholes or insufficient safeguards in policymaking, such as assisting corruption prevention officers in evaluating legislation to assess corruption risk factors in legal texts in Lithuania. AI can also support governments' connection with the public to reinforce accountability. This is the case of some online platforms and tools, the virtual assistant Chatico from Bogotá (Colombia), which has an open government module that eases participation in public campaigns and decision-making processes.

Table 2.5. Examples of AI for enhanced accountability and anomaly detection

Country	Initiative	Description	Sub-benefit	Function
Chile	ChileCompra	ChileCompra's Public Contracting Observatory uses LLM's to analyse procurement data for irregularities and improve compliance monitoring, enabling more efficient oversight and promoting ethical standards in public procurement.	Detecting improper transactions and assessing integrity risks	Public procurement (Box. 5.24)
Colombia	Chatico	The city of Bogotá launched Chatico, an AI-powered virtual assistant, to facilitate interactions between citizens and the local administration. Through its website and WhatsApp interfaces, the	Enabling non-governmental actors to understand and	Civic participation (Box 5.41)

		chatbot eases citizen participation in public campaigns and decision-making processes and offers as well enhanced accessibility to public services.	engage with government and promote accountability	
European Union	DATACROS	DATACROS uses AI to detect anomalies in corporate ownership structures that may indicate corruption or money laundering. The system analyses data from over 70 million companies across 44 European countries, flagging hidden patterns and potential illicit activities.	Detecting improper transactions and assessing integrity risks	Fighting corruption and promoting integrity (Box 5.27)
Lithuania	AI to assist corruption prevention officers	Lithuania is developing an AI-powered tool that uses large language models (LLMs) to assist corruption prevention officers in evaluating corruption risk factors in legal texts, such as loopholes or insufficient safeguards.	Detecting improper transactions and assessing integrity risks	Fighting corruption and promoting integrity (Box 5.30)
Portugal	Assessing procurement risks	Portugal Court of Audit is implementing AI-driven risk assessment methods to enhance its audits, ensuring the most critical cases in public procurement receive priority. This initiative optimises resources and strengthens accountability in public contracting.	Detecting improper transactions and assessing integrity risks	Public procurement

Unlocking opportunities for external stakeholders through AI as a good for all

Finally, a small minority of use cases (4%) have the potential to unlock opportunities for external stakeholders, such as citizens, civil society organisations and businesses. Such activities are different from general service delivery or new forms and channels of interaction for participation and accountability purposes. Here, AI can empower external actors by enhancing their capabilities and access to information and use government-supported AI systems to achieve their missions and objectives more effectively. Such uses remain marginal and represent a potential gap in government efforts warranting further research and action. As noted above, however, such use cases could potentially be more prevalent in areas not covered by this report.

Table 2.6 provides some examples of how AI is being used for these purposes. Use cases generally include tools in participatory platforms that can be used according to the needs and objectives of users. For example, the participatory platform Decide Madrid (Spain) experimented with AI tools to assist citizens in aggregating and developing their own proposals for action. Similarly, the AI tool MAPLE helps citizens by summarising draft legal texts and allowing them to submit inputs and comments on pending legislation. In the field of disaster risk management, governments can also contribute to goods that can empower citizens for greater resiliency to natural disasters. For example, the Bencana Bot in Indonesia prompts residents to report floods via social media, generating real-time online maps that, combined with official data, can be reused by non-governmental stakeholders. Although the examples here help address the relevant benefit, they do so in a somewhat tangential way, with external elements sometimes benefiting as a positive spillover effect from government activities. Emergent opportunities for governments to open or provision AI systems more directly for external stakeholders may result, such as in areas not currently serviced or appealing to the private sector.

Table 2.6. Examples of AI unlocking opportunities for external stakeholders through AI as a good for all

Country	Initiative	Description	Function
Greece	DidaktorikaAI	Greece's DidaktorikaAI platform, launched by the National Documentation Centre (EKT), improves the accessibility of academic and scientific knowledge for policymakers and the broader society through an AI-powered online library gathering more than 50 000 publications.	Civic participation and open government
Indonesia	Bencana Bot	In Jakarta (Indonesia), the AI-powered chatbot Bencana Bot prompts residents to report floods via social media, generating real-time, freely accessible online maps on PetaBencana.id. Combined with official data provided by the Jakarta Disaster	Law enforcement and disaster

		Mitigation Agency, the platform helps residents stay safe during emergencies, with usage surging by 2 000% during major floods.	risk management
Spain	Decide Madrid	In 2021, the participatory platform Decide Madrid (Spain), based on the open-source software Consul, experimented with a Natural Language Processing (NLP) system to assist citizens in aggregating and developing proposals.	Civic participation and open government
United States	MAPLE	The AI-powered tool MAPLE (Massachusetts Platform for Legislative Engagement), which allows citizens to better understand the context and objectives of draft legal texts through AI-generated summaries and to submit their inputs and comments to pending legislation.	Civic participation and open government

Some government functions are more mature regarding governing and adopting AI

AI's potential is recognised across all functions of government, but the maturity of its adoption and governance varies significantly. A few functions are already implementing AI initiatives in a structured manner, learning from their implementation, and in some instances, scaling up successful solutions into other issue areas and in broader contexts. For instance, in public service design and delivery, AI is widely used to automate tasks, retrieve and synthesise information, and improve digital service effectiveness and usefulness for users. In law enforcement and disaster risk management, AI is widely used to prioritise police resources, accelerate investigations and better anticipate and recover from disasters. However, other functions of government have only begun experimenting with limited, ad-hoc prototypes and pilots. In fields such as policy evaluation, public financial management, and regulatory design and delivery, AI adoption is largely found in isolated pilots. In some functions, such as justice administration and fighting corruption and promoting public integrity, there is significant geographical variation. Some countries, like Argentina (Box 5.62 on Prometea) and Spain (Box 5.67 on AI-enabled domestic violence response), actively deploy sophisticated AI solutions in even high-risk areas that manage to mitigate risks otherwise resulting in scandals with similar systems in other contexts. Yet, other countries are still in the early stages of digital transformation. As AI continues to appear in additional use cases, governments must mitigate risk in order to promote responsible AI adoption.

In terms of technical maturity, AI adoption varies not only in scale but also in the types of systems used. Some functions of government, such as tax administration and procurement, rely heavily on rules-based systems, which have been effective in automating structured decision-making processes for many years. Others, like law enforcement and disaster management, and fighting corruption and promoting public integrity, use more advanced ML systems to identify patterns, enhance risk assessments and support decision-making. However, in most functions of government, there is very little use of the latest GenAI models, such as LLMs, which offer new capabilities in knowledge synthesis and content generation that could be more transformational. This trend can be seen in other databases as well — only 61 of the 1 343 (4.5%) AI use cases in the EC Public Sector Tech Watch repository are GenAI (Brizuela et al., 2025[16]). A survey from Deloitte (2024[17]) also found an uneven level of preparedness for GenAI adoption across different sectors of government, though their sector categorisations do not align directly to the government functions in this report. The use of GenAI systems can be seen in a handful of government functions in Chapter 5, such as the design and delivery of regulations and public services, and in civic participation efforts, though many of these efforts appear sporadic or experimental. Some of the most advanced uses take the form of chatbots, which may be highly impactful but may not fully exploit the technology's potential for large-scale synthesis, tailored content generation or making services more proactive or personalised. The slow adoption of these advanced systems in many domains, in addition to the relative recency of the technology and relevant AI use cases, suggests that while AI experimentation is widespread, the transition to more sophisticated, high-impact systems remains uneven across government functions and countries. This is not to say that governments should abandon all efforts in using more established forms of AI in pursuit of GenAI, as chasing the latest "cool tech" in areas where it is not a good match for the problem to be solved can contribute to AI project failure (Ryseff and Narayanan, 2025[18]).

A key factor shaping AI adoption across functions of government is data availability and quality. In tax administration, for example, AI has been widely deployed due to the abundance of structured data, which has enabled automation and risk assessment for years. Yet, because of the complex legal landscape in this function and regarding taxpayer data, most efforts rely on classic rules-based systems, with challenges in pursuing more modern ML systems that could unlock productivity gains even leveraging unstructured data. By contrast, while AI is increasingly used in private sector HRM functions, its application in civil service reform remains limited due to the insufficiency of comprehensive workforce data — covering employee skills, job demands and performance indicators. The abundance of data in governments does not necessarily translate to the availability of AI-ready data (see Chapter 3 on implementation challenges).

Beyond data, other fields could be facing challenges in AI maturity because the nature of their tasks requires technical capabilities that only recent advanced AI systems could provide. This suggests the potential for rapid acceleration in the coming months. For instance, AI adoption in regulatory design and delivery appears to have increased and accelerated through the use of LLMs to assist with analytical tasks that could have not been performed by other systems (for example, see Box 5.12 for legal and regulatory querying and drafting aids). In another example, AI is enabling mass deliberative civic participation efforts at scales never before feasible.

Different functions of government have different contexts and needs

AI adoption in government is often discussed in broad terms, but its impact and risks vary significantly across functions. Different functions have unique challenges, regulatory constraints and levels of AI readiness. For example, when AI is used in public services to improve healthcare, the applications need to navigate stringent data privacy regulations and ethical concerns around medical decision-making, while AI in other fields, such as civic participation, can be more experimental, using real-time optimisation with less potential to infringe the rights of individuals. Actors in one function of government could use a similar AI solution as in another with vastly different results, impacts and implications. Thus, the discussion of function of government in this section cannot be seen as a likewise comparison, and further research is needed to understand differences, including with a larger scope of analysis.

Use cases could pose risks if not implemented in a trustworthy manner

As discussed in Chapter 1, this report categorises five types of risks faced by governments in adopting AI. In addition to the risks shown in Figure 2.7, the risk of inaction involves missed opportunities and the growing capacity gap between the public and private sector. Such risks, possibly realised by *not* building capacities for and using AI in government, cannot be visualised and may be difficult or impossible to precisely measure.

The OECD found that every one of the 200 use cases analysed for this report could pose one or more types of risk if not designed and used in a trustworthy way.

Figure 2.7. The potential for operational risks is the most represented across government functions

Number of use cases across functions of government categorised according to selected risk types
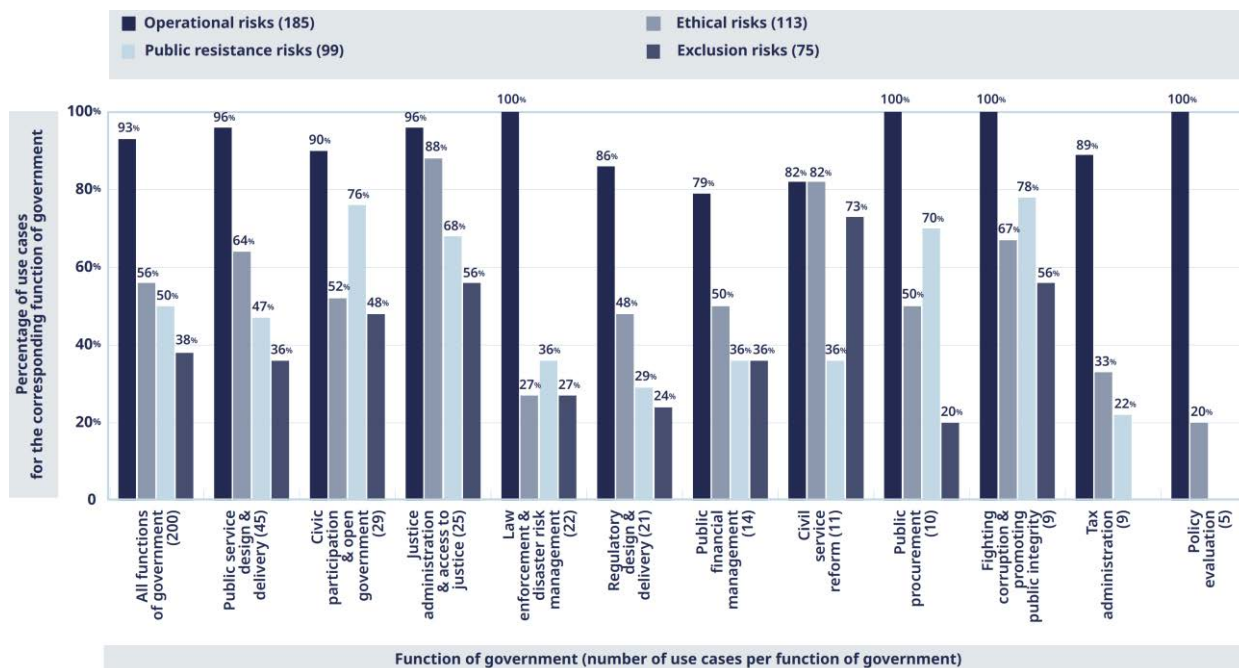
Note: In parentheses, the number of occurrences of risk types. Use cases can involve more than one type of risk. Thus, the total number of potential risk occurrences is greater than the total number of use cases.
Source: OECD analysis of identified use cases.

Potential **operational risks are the most prevalent** among the analysed used cases (93%). An example of this is Australia's Robodebt scheme, where investigations revealed inadequacies in algorithmic design played a significant role in its failures and its calculations ultimately being ruled unlawful (Box 5.11). Specifically, the algorithm's oversimplification and lack of safeguards resulted in the issuance of 470 000 incorrect debt notices without human verification. This highlights the operational risks of automating complex social systems without sufficient human oversight or rigorous testing and additionally presented ethical risks resulting in real-world harm.[4]

Potential **ethical risks were the second most prevalent** among the analysed use cases (56%), such as in AI-supported applicant review tools used by HRM offices. In reality, ethical risks can be present in the large majority of AI uses cases. However, the lower presence of this risk in the analysed use cases could be due to the limited and specific scope, and manner of application, of use cases identified for this report. An example of an ethical risk that resulted in real-world harm is The Netherlands' Toeslagenaffaire (childcare benefits scandal), where an AI system wrongfully accused 26 000 families of fraudulently claiming childcare benefits due to a biased algorithm that targeted families with dual nationalities or migrant backgrounds; it forced many to repay undue debts (Box 5.6). This case illustrates how ethical risks can cause harm if not mitigated appropriately.

Potential **public resistance risks were prevalent across 50% of the analysed use cases**. Previous failures in AI deployment have significantly impacted reputations and eroded public trust in the government's capacity to use AI responsibly. These cases underscore the necessity for governments to take steps to prevent risks and to swiftly address potential failures in AI use, fostering public trust. This can be achieved with appropriate guardrails, including strong accountability and redress mechanisms, continuous monitoring and oversight, and effective risk management.

Finally, the potential for **risks of exclusion was identified in 38% of the analysed used cases**. For example, citizen participation platforms that have deployed AI tools to assist citizens in aggregating and developing proposals may pose challenges for individuals lacking digital skills. This could inadvertently benefit the advantaged, thereby enhancing their ability to further promote their ideas (Duberry et al., 2021[19]; Wang et al., 2024[20]). AI use cases also help civil servants to efficiently process vast quantities of citizen inputs and enhance facilitation of participatory processes. However, the success of these applications is contingent upon the diversity incorporated in the training data of the AI system employed; there is a risk these tools may fail to adequately capture the diversity of public opinion (ECNL, 2024[21]). Many AI-assisted translation tools in participation platforms may also be subject to not capturing nuances and understanding different cultural contexts for languages of minorities, as they are usually trained on data from English and other dominant languages (ECNL, 2024[21]).[5]

The greater or lower presence of certain risks across the functions of government is of interest, too. For instance, observed cases in justice administration and civil service reform appear to more prominently feature ethical risks due to the potential for adverse outcomes in some of its cases. In the tax administration field, existing controls and regulations might create safeguards against ethical and exclusion risks, making operational risks more prevalent in the field. In civic participation, public resistance risks take more prevalence, compared to other functions. This is likely because most of its use cases include government-to-citizen interactions and thus exhibit greater dependence on public acceptance. This is similar for public procurement, where suppliers' perceptions of and public trust in AI systems play a key role in their success; thus featured more prevalently in resistance risks. This comparative analysis shows the importance of acknowledging the main drivers of potential risks in each field, which can inform how to mitigate and manage them.

Governments are vigilant of several AI risks, though some may be overlooked

Overall, through conducting analysis and interacting with governments in the development of Chapters 4 and 5, it appears many national governments are well-informed with regard to, and have put in place, processes to manage risk associated with **data, a lack of transparency and explainability, and AI misuse** — either intentional or inadvertent — and the potential for resulting harms or privacy infringement. This is positive and not surprising, as these risks are often raised by AI experts and in research both within and beyond the public sector (OECD, 2024[13]).[6] To a somewhat lesser extent, the risk of an overreliance on AI technologies also appears to be a consideration for government efforts. Overall, however, there appears to be less of an emphasis on ensuring that government use of AI does not further exacerbate digital divides. When complementary service channels are not made available, government ambitions to automate and streamline processes could result in the reduction of service opportunities for communities with less access to digital services or preference for non-digital approaches (Welby and Hui Yan Tan, 2022[22]). In addition, while governments are clearly aware of how AI could contribute to productivity and help shift public servants' efforts away from repetitive tasks and towards more meaningful work, there is seemingly less recognition of AI's potential to reduce job quality (e.g. through invasive algorithmic management) or for job displacement (Peixoto, Canuto and Jordan, 2024[23]). Even if AI adoption becomes systematic in public administrations, governments will need to ensure that all citizens are properly served and the concerns and any rights of the civil service are taken into account in order to promote AI adoption.

Finally, governments need to better consider the **risk of inaction**. Throughout Chapters 4 and 5, it is clear that many governments are working towards specific goals with their AI efforts, and that they are aware of and are seeking to mitigate a variety of AI risks by instantiating approaches to unlock the potential of AI. However, the analysis conducted for this report and discussions with relevant government officials indicate there may be limited awareness of missed opportunities due to slow AI adoption or the consequences of the widening gap in AI capabilities between the public and private sectors.

Few governments have assessed the extent to which AI could make an impact in their internal operations and public-facing services. Furthermore, most have not fully articulated their ambitions for AI in government, determined the existing gaps, or proposed clarified roadmaps to close those gaps and meet the goals. Governments need to explore AI not only to enhance the design and implementation of public policies and services, but also to ensure they have the knowledge and capacity to regulate AI development and use in government and beyond. In an inquiry by the Parliament of Australia (2025[24]), the Joint Committee of Public Accounts and Audit expressed "very grave concern" that AI will soon outpace the government's ability to regulate it. Such committee findings are not necessarily representative of the Australian Government's views. AI experts have also noted the inability of governance mechanisms and institutions to keep pace with rapid AI evolutions is one of the most critical risks associated with AI (OECD, 2024[13]).

In addition to these critical risks associated with the AI use in government, the OECD's analysis of 200 use cases has identified a variety of challenges governments can face in adopting the technology. These implementation challenges are shared across all functions, while others are more prevalent in certain fields. They can translate into broader issues and can hinder the strategic use of AI in government. These issues are discussed in the following chapter.

References

Austin, T. et al. (2024), *A snapshot of how public sector leaders feel about generative AI*, https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-adoption-in-public-sector.html. [17]

Bailey, K. (2016), *Reframing the "AI Effect"*, https://medium.com/@katherinebailey/reframing-the-ai-effect-c445f87ea98b. [12]

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. [15]

Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, European Commission, https://op.europa.eu/s/z4XY. [16]

Columbia University (2020), *The Future of AI in the Brazilian Judicial System*, https://www.sipa.columbia.edu/aidriven-innovations-brazilian-judiciary. [9]

Duberry, J. et al. (2021), "Artificial intelligence and civil society participation in policy-making processes: Thinking about AI and participation.", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3817666. [19]

EC (2025), *Public Sector Tech Watch latest dataset of selected cases*, http://data.europa.eu/89h/e8e7bddd-8510-4936-9fa6-7e1b399cbd92 (accessed on 4 April 2025). [10]

ECNL (2024), *Can AI tools and platforms make public engagement truly meaningful and inclusive?*, https://ecnl.org/news/ai-public-participation-hope-or-hype (accessed on 17 March 2025). [21]

Government of Canada (2025), *AI Strategy for the Federal Public Service*, https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/gc-ai-strategy-overview.html. [4]

Government of Switzerland (2025), *Strategy Use of AI systems in the Federal Administration*, https://www.bk.admin.ch/bk/en/home/digitale-transformation-ikt-lenkung/ikt-vorgaben/strategien-teilstrategien/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html. [5]

Government of Uruguay (2021), *AI Strategy for the Digital Government*, https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/publicaciones/ia-strategy-english-version/ia-strategy-english-version/ai-strategy-for. [6]

Harvard Kennedy School (2023), *AI, judges and judgement: setting the scene*, https://www.hks.harvard.edu/centers/mrcbg/publications/awp/awp220. [8]

Muñoz-Cadena, S. et al. (2025), *Sistemas de IA en el sector público de América Latina y el Caribe (Versión V2)*, https://sistemaspublicos.tech/sistemas-de-ia-en-america-latina/ (accessed on 29 April 2025). [11]

Netherlands Court of Audit (2024), *Central government often does not assess risks of AI*, https://english.rekenkamer.nl/latest/news/2024/10/16/central-government-often-does-not-assess-risks-of-ai. [25]

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. [7]

OECD (2024), "Assessing potential future artificial intelligence risks, benefits and policy imperatives"*, OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, https://doi.org/10.1787/3f4e3dfb-en. [13]

OECD (2024), "Governing with Artificial Intelligence: Are governments ready?"*, OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, https://doi.org/10.1787/26324bc2-en. [1]

OECD (2023), "The state of implementation of the OECD AI Principles four years on"*, OECD Artificial Intelligence Papers*, No. 3, OECD Publishing, Paris, https://doi.org/10.1787/835641c9-en. [2]

OECD (2021), "State of implementation of the OECD AI Principles: Insights from national AI policies"*, OECD Digital Economy Papers*, No. 311, OECD Publishing, Paris, https://doi.org/10.1787/1cd40c44-en. [3]

Parliament of Australia (2025), *Report 510: Inquiry into the use and governance of artificial intelligence systems by public sector entities - 'Proceed with Caution'*, https://parlinfo.aph.gov.au/parlInfo/download/committees/reportjnt/RB000567/toc_pdf/Report510Inquiryintotheuseandgovernanceofartificialintelligencesystemsbypublicsectorentities-'ProceedwithCaution'.pdf. [24]

Peixoto, T., O. Canuto and L. Jordan (2024), "AI and the Future of Government: Unexpected Effects and Critical Challenges", *Policy briefs on Economic Trends and Policies*, Vol. 2408, https://ideas.repec.org/p/ocp/pbecon/pb_10-24.html. [23]

Ryseff, J. and A. Narayanan (2025), *Why AI Projects Fail*, https://www.rand.org/pubs/presentations/PTA2680-1.html. [18]

The Alan Turing Institute (2024), *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions*, https://www.turing.ac.uk/news/publications/ai-bureaucratic-productivity-measuring-potential-ai-help-automate-143-million-uk. [14]

Wang, C. et al. (2024), "The artificial intelligence divide: Who is the most vulnerable?", *New Media & Society*, https://doi.org/10.1177/14614448241232345. [20]

Welby, B. and E. Hui Yan Tan (2022), "Designing and delivering public services in the digital age", *OECD Going Digital Toolkit Notes*, No. 22, OECD Publishing, Paris, https://doi.org/10.1787/e056ef99-en. [22]

Yigitcanlar, T. et al. (2024), *Local governments are using AI without clear rules or policies, and the public has no idea*, https://theconversation.com/local-governments-are-using-ai-without-clear-rules-or-policies-and-the-public-has-no-idea-244647. [26]

Notes

[1] Not all of the use cases analysed appear in this report. The OECD made a selection of cases for the report that best illustrated various themes and findings presented.

[2] See https://oecd-opsi.org/innovation-tag/artificial-intelligence-ai and https://oecd.ai/dashboards/policy-instruments/AI_use_cases_in_the_public_sector, respectively. The OPSI collection included a global open "Call for Innovations" crowdsourcing exercise focused on innovations in public services in 2024.

[3] As categorised in the observatory's *Classification of the Functions of Government*. See https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary%3AClassification_of_the_functions_of_government_%28COFOG%29.

[4] The Robodebt scheme leveraged automated data-matching, income averaging and overpayment calculation. As discussed in Box 1.1 of this report, many argue that such systems should not be considered AI at all. Thus, the Robodebt scheme could be better described as an automated decision-making system. Nevertheless, it helps to illustrate issues in governance, ethical oversight and algorithmic design.

[5] See also the related discussion on Exclusion Risks in Chapter 1.

[6] While governments may be informed on these and other issues, and as indicated in Chapter 4, many have frameworks and processes in place to mitigate risks, they need to take action and follow through on their frameworks and processes in order to manage risks. There are instances in which governments may not fully comply with internal requirements or may face incentives to consider systems as low-risk, thus reducing accountability requirements (Netherlands Court of Audit, 2024[25]). In addition, some research suggests that local governments may not be as informed or prone to undergoing risk mitigation activities as national governments (Yigitcanlar et al., 2024[26]). Finally, this report does not seek to evaluate the quality and effectiveness of government processes and mechanisms, though it does seek to highlight those that appear sound and are emerging as best practices.

3 Implementation challenges that hinder the strategic use of AI in government

This chapter examines challenges seen in analysed use cases and broader research and analysis. It finds that many government AI initiatives remain in pilots. It highlights common system-wide barriers – skills gaps, difficulties accessing and sharing high-quality data, limited actionable guidance, risk aversion, and weak measurement of results and return on investment. Other challenges vary by function, including inflexible or ambiguous regulation, high or uncertain costs, and outdated legacy information technology systems.

Key messages

- The OECD's review of AI use cases in government indicates a high presence of early-stage initiatives, such as experiments and pilots. This indicates:
  - o Possible challenges in transitioning from experimentation to implementation
  - o A need for increased monitoring and sharing of information
  - o A possible need for policy actions to encourage implementation and scaling.
- Implementation challenges that are shared across all government functions are:
  - o Skills gaps
  - o Challenges to obtaining or sharing quality data
  - o A lack of actionable frameworks and guidance on AI usage, including for specific policy areas
  - o Risk aversion
  - o Demonstrating results and return on investment (ROI).
- Government functions face a variety of other challenges, with some more prevalent in some functions than others:
  - o Inflexible or outdated legal and regulatory environments
  - o High or uncertain costs of AI adoption and scaling
  - o Outdated legacy information technology (IT) systems.

Governments face a variety of implementation challenges in adopting AI. Some of these challenges span all functions of government, while others appear more acute in certain ones. For instance, nearly all areas struggle with skills gaps and accessing and sharing quality data. AI implementation should account for distinct regulatory landscapes in different functions, with compliance requirements varying significantly between functions such as public procurement, law enforcement and tax administration. Additionally, some functions are more challenged than others with securing funding or with outdated systems, as only now are we seeing the emergence of digital infrastructure that share services with greater interoperability and integration. These challenges can translate into broader issues, such as difficulties in scaling up solutions and risk aversion that hinders innovation. This chapter discusses the various implementation challenges faced by functions of government, as further discussed as related to each function in Chapter 5. Many of these challenges mirror specific regional analysis on AI in government (OECD/CAF, 2022[1]; Brizuela et al., 2025[2]).

Most government AI efforts exist in exploratory or pilot phases, with limited scaling and documentation

Possible challenges in moving from experimentation to implementation

The OECD's review of AI uses in government indicate a high presence of early-stage initiatives, such as experiments and pilots. This is consistent with discussions held among relevant OECD working parties and networks, in which government officials at both national and local levels report being in the early stages of using AI government, seeking to learn by starting small and testing different approaches.

Overall, this approach is a good one. The OECD has long encouraged governments at both national and local levels to experiment with new approaches in a controlled and iterative manner to minimise risks and

costs and to promote failures — which are inevitable — to occur quickly and generate lessons learned for future efforts (OECD, 2017[3]; 2024[4]). This is critical, especially when first getting used to using AI, as some estimates suggest that more than 80% of AI projects fail, double the rate of non-AI projects (Ryseff, De Bruhl and Newberry, 2024[5]). For instance, a successful path to AI adoption can involve incorporating AI into low-risk areas and processes and using internally-generated or open data to demonstrate value and establish quick wins. An example of this case is the Finnish Government Shared Services Centre for Finance and HR (Palkeet), which began with modest applications of RPA, paving the way towards "hyper-automation" with machine learning (ML) (see Chapter 5, "AI in public financial management").

Yet the end goal of most AI projects is the eventual implementation, and as appropriate, the scaling up of successful solutions. In many countries and functions of government, AI use cases largely exist in the exploratory phase (e.g. proofs of concept, pilot projects) and have not yet been more broadly implemented or scaled beyond limited use. For instance, in the United Kingdom (UK) — generally one of the more advanced governments when it comes to using AI — a report by the Parliament's Public Accounts Committee (PAC) (2025[6]) found the government had "no systematic mechanism for bringing together learning from pilots and there are few examples of successful at-scale adoption across government". The results of a Deloitte survey (2024[7]) in 14 countries reinforce challenges in scaling GenAI in particular, noting also this challenge is not unique to government. According to the survey, "a significant number of both commercial and government respondents have transitioned fewer than one-third of their GenAI experiments into full production". This is attributed to other challenges discussed in this chapter: lack of expertise and difficulty in measuring mission value from GenAI.

Existing data sources provide insights about the current state of AI adoption in government. For example, the use cases analysed for this report show most have moved beyond small pilots to be more fully implemented in some way. Of these implemented cases, most have not scaled beyond their original contexts (e.g. in certain offices or for certain processes) to address other needs. Although this is not always a goal, and AI uses in some government functions are often not appropriate for others. AI use cases in public service design and delivery in Chapter 5 demonstrate, however, that successful approaches can indeed be scaled up. In this report, the preponderance of implemented cases that go beyond pilots is in part due to a tendency for the OECD to select more implemented use cases because there is more public information available about them, and governments are somewhat more likely to report on them as part of OECD data collection exercises.

Other data sources offer complementary perspectives. The Public Sector Tech Watch observatory of the European Commission (EC) has systematically collected AI uses cases for several years. Its data on nearly 1 500 AI use cases indicate that most AI solutions are still in the planned, pilot or in-development phase (58%), suggesting across the EU public sector the majority of cases remain experimental or not fully implemented (Figure 3.1). Although moving from pilots to production appears to be a challenge — as reinforced in OECD discussions with governments — the proportion of implemented projects has increased in the latest data collections. This suggests that administrations may be transitioning their initiatives from initial testing to fuller implementation (EC, 2024[8]). This data does not consider the extent to which implemented projects have scaled beyond their initial context.

Figure 3.1. Most European Union (EU) AI use cases are in pilot or development phases
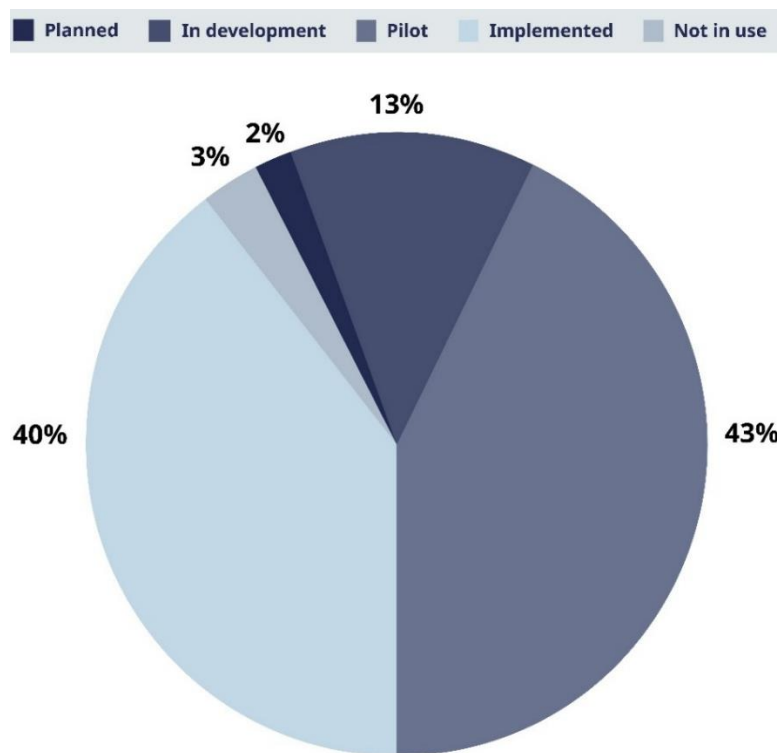
Source: Data analysed by the OECD from (EC, 2024[9]).

The "AI Systems in the Public Sector in Latin America and the Caribbean" database indicates a much higher proportion of implemented use cases (70%) (Muñoz-Cadena et al., 2025[10]). However, this is likely due to the data collection for this database, which was built by researchers using publicly available information. Publicly available information may be less likely to include details on planned or piloted use cases compared to submissions from governments, which largely informed the OECD and EU databases.

Overall, based on the OECD's work on AI in government since 2019, including directly with many governments, it seems the expected state of implementation should be higher relative to early-stage testing. The high presence of early-stage cases would be easily explained if government use cases had a heavy emphasis on leveraging the latest AI technologies, such as generative AI foundation models. However, OECD research for this report indicates this is largely not the case; governments still tend toward more longstanding approaches. These observations suggest further work in this area is warranted to gain additional understanding about governments planning-to-implementation and scale-up journeys, and to derive more specific lessons learned and factors for success.

As discussed in Chapter 2, it may be possible that some narrow but traditional applications of AI — which may be more likely to be implemented or even scaled up — may have become so integrated or commonplace that they no longer trigger external reporting or a response to data collection efforts. This could skew the numbers through underrepresentation of implemented initiatives that do not rely on newer AI systems and approaches.

Lack of evidence on continuity calls for increased monitoring

Additional information is required to gain a better understanding of the status and evolution of government AI use cases. Future research should investigate this further by monitoring the progression of solution development over time, which could generate lessons from both successes and failures. The OECD sought

to explore the progress in relevant AI use cases discussed in previous products; however, there are generally few updated reports on the results of pilots or the status of implemented use cases, making it challenging to follow them in a longitudinal way. Most of the research done to determine whether a use case is operational relies on the ability to access a functional product (if public), or the presence of government press releases, news coverage, blog articles or public presentations. Another source is use cases shared through periodic public reports or official repositories, including those by the OECD. Current inventories or catalogues are often static; they depict projects as a snapshot in time, without providing insights into their development and evolution, and are often not updated. Primary data collection by researchers (i.e. directly surveying government organisations to further identify new efforts and obtain updates and lessons learned on known initiatives) would be useful for further research, though demands greater resources and time.

These challenges in accessing up-to-date information on individual use cases underscores the importance of governments' monitoring of AI use cases, along with a systematic and regular sharing of information, which would further reinforce the OECD AI Principles on transparency and accountability. This would not only be valuable for external audiences; documenting and disseminating successful (and unsuccessful) methods and use cases can help government organisations to replicate and scale AI projects more effectively. This approach helps avoid common errors, helps ensure consistency and accelerates the adoption of AI technologies across various government entities (OECD/UNESCO, 2024[11]). Inadequate or absent monitoring can also affect future use cases, as potentially effective AI innovations may be overlooked or, conversely, disproven approaches may be scaled up inappropriately. Further discussion and examples of how some governments are doing this can be found in Chapter 4, "Promoting transparency in how government uses AI".

Policy actions may be needed to encourage implementation and scaling

The fact that many AI initiatives are in the planning and pilot phases, or are unclear regarding their progression, suggests governments need to enhance their implementation capabilities to advance projects beyond initial testing stages, secure successful deployment and sustain long-term impact (EC, 2024[8]). This involves establishing foundational elements such as ensuring access to datasets, computing resources and the necessary expertise required to develop and scale AI projects. Such factors are discussed in-depth in Chapter 4. It also requires overcoming other implementation challenges, as discussed below.

Common challenges shared across core government functions

Skills gaps, the most common challenge

A recent survey in five countries from Salesforce (2024[12]) found a lack of internal skills for using AI to be the primary barrier to government AI adoption, with 60% of public sector respondents highlighting this challenge.[1] Public sector respondents were a third more likely to indicate a skills gap in their organisation compared to the industry average. National-level reviews find comparable results, with 70% of UK government bodies reporting skills as a barrier to AI adoption (UK NAO, 2024[13]). In a National Trade Union Study of 2 000 Australian Public Service (APS) employees from August to October 2024, 92% said they had received no training on using AI, and only 16% said they felt equipped to use the technology.[2] The Australian government has released and made available to all APS employees an AI in Government Fundamentals training module (in October 2024) and a series of MasterClass sessions on AI run by practitioners. The Australian Government has a number of capability building structures in place to enhance AI capability. The AI CoLab initiative provides a framework for cross-sector collaboration,

codesign and regular events. Access to a government AI tool through a closed Beta Trial of GovAI commenced on 5 May 2025 with expanded access available to all APS employees on 31 August 2025.

Skills gaps are also a significant challenge when focusing specifically on sub-national governments (UN Habitat, 2024[14]). For instance, United States (US) surveys of state Chief Information Officers and local-level IT executives found that only 20% of state CIOs and 25% of local respondents were even slightly confident their technology workforce possessed the expertise necessary to respond to the advent of generative AI (NASCIO, 2024[15]; PTI, 2024[16]). Beyond the commonly discussed challenge of governments competing with the private sector for talent, in some instances, national and sub-national governments are in competition with one another for the same limited talent pool. Smaller cities can also suffer from "brain drain", with young talent moving to larger cities that provide more career possibilities (de Mello and Ter-Minassian, 2020[17]).

Skills gap challenges can be seen across nearly all functions of government discussed in Chapter 5. Skills challenges limit governments' ability to take advantage of the latest developments in AI and can contribute to reluctance among public servants to accept the use of AI in general. In several functions, governments are struggling to determine exactly what kinds of skills are needed, and for whom.

Skills gaps can exacerbate other risks and challenges (Trajkovski, 2024[18]). For instance, they can lead to poor outcomes, the overestimation and misplaced trust in AI capabilities and systems, inadvertent misuse and general non-compliance with laws and other rules. Inconsistent levels of skill in governments can lead to pockets of innovation, but with little ability to scale beyond them.

In addition, a lack of skills internal in public administrations can result in an overreliance on outsourcing through public procurement (Mitchell, 2025[19]; Autio, Communigs and Elliott, 2023[20]). While procurement is an important and normal aspect of obtaining AI-related goods and services, relying too heavily on procurement relative to building internal capacities can result in a hollowing-out of government capacities (Trajkovski, 2024[18]). This can form a vicious cycle, where governments lack the right skills to design upskilling programmes, understand which skills to recruit for, and fully understand vendor offerings to procure the right goods and services at a fair price. Overall, without proactive skills development, "public agencies will find themselves merely reacting to technological shifts rather than steering these emerging technologies to serve societal interests effectively" (Trajkovski, 2024[18]). If government cannot demonstrate efficient and effective use of and self-control of AI, then it is unlikely that it will be able to regulate the technology. This also contributes to the challenge of high costs for AI adoption, as discussed below; hiring contractors can cost three to four times as much per person as government employees (UK DSIT, 2025[21]).

Several governments have instantiated upskilling and targeted recruitment programmes (see Chapter 4, "Fostering skills and talent"), with some even using AI as a tool to achieve these goals (see Chapter 5, "AI in civil service reform").

Lack of high-quality data and the ability to share it

Through all levels of government and nearly every government function discussed in-depth in Chapter 5, data challenges are an impediment to developing and using AI in government. Recent work by RAND (2024[5]; 2025[22]) found that data issues were one of the main drivers in failed AI projects, including a lack of suitable data, and noted the importance of work often perceived to have low "activity prestige" (e.g. data cleaning).

For some functions, the needed data may simply not exist or were never digitised from paper (e.g. as seen often in justice administration), or the quality of data available is deficient (e.g. poorly structured, incomplete or mistyped records, discrepancies in data formats). This can arise from a variety of reasons, including poorly controlled data input processes, or even previous lack of foresight that such data could someday

be important. While time-consuming and burdensome, such quality issues can often be overcome, such as through digitisation, data cleaning and validation processes.

The repeated emphasis on a lack of sufficient, quality data may seem counterintuitive, as governments hold tremendous amounts of data and often make it available as open government data (OGD) to, among other things, help serve as inputs for training AI systems (OECD, 2023[23]). However, in many cases, it is the ability for government agencies to share data amongst themselves that is a challenge. This can be due to rules that prevent sharing or are unclear to public servants, months-long approval processes, or less commonly, governments signing away data reuse rights in contracts with companies. These issues are also reflected when sharing data between jurisdictions and levels of government, adding challenges for AI adoption in subnational governments. Some countries are grappling with how to comply with data protection and management rules, such as the EU General Data Protection Regulation (GDPR), though these data management issues generally existed long before such rules were put into force. In other cases, there is a lack of technical or policy protocols for sharing, or lack of interoperability across IT systems and data formats. In most cases, these issues are symptoms of a more systemic problem — inadequate data governance resulting in non-strategic, sporadic and fragmented data collection and management — and associated rules across government. Antiquated or burdensome rules around data sharing also contribute and may need to be reconsidered to account for technological advancements while still protecting privacy.

Overall, it is critical that governments establish sound data governance and management activities to succeed in adopting AI — although only 59% of OECD countries have a data strategy in place for the public sector, with even fewer providing actionable guidance for implementation (OECD, 2024[24]). Without strong data governance in place, governments risk developing and deploying AI systems that use poor quality data, resulting in outcomes ranging from simple inaccuracies to systemic bias and unfair outcomes for citizens. Without robust data governance across organisations and levels of government, governments' AI ambitions would largely need to be limited to the small experiments and pilots in limited settings that are the norm today, as discussed above.

Fostering the development of data-driven public sectors has long been a focus of the OECD (2019[25]; 2023[23]), and a number of governments are putting in place the data foundations needed for governments to reap AI's benefits (see Chapter 4, "Creating a strong data foundation").

Lack of actionable frameworks and guidance on AI usage

National strategies for AI in government — either dedicated strategies or those embedded in broader instruments — are now common, and they are important in defining a vision for AI success. However, they generally provide only high-level details on commitments and aspirations, offering limited concrete guidance to facilitate the materialisation of AI's benefits while safeguarding against its risks. They also often fail to address key operational considerations that would make them effective. Investments and procurement, for instance, are often overlooked, despite being crucial for AI in government (van Noordt, Medaglia and Tangi, 2023[26]; Monteiro, Hlacs and Boéchat, 2024[27]). To bridge this gap, governments need actionable guidance that is aligned to strategies and provides their institutions with tangible direction and assurances. Guidance is also important for sub-national governments, such as cities. This can be important for both ensuring alignment with national approaches — as sub-national governments often follow or take inspiration from national efforts — as well as for helping sub-national governments meet their own digital and AI ambitions as well as the needs of their citizens and residents.

Such guidance can be either boundary spanning, addressing system-wide issues to promote trustworthy AI adoption, or targeted, focusing on specific government functions and applications. Cross-cutting guidance provides clarity and direction for fundamental elements such as data governance, talent development and investment. Vertical guidance, in contrast, helps public servants navigate AI's opportunities and risks via means effectively tailored to different policy domains. These approaches are not exclusive and can both be pursued complementarily.

Overall, there is a lack of concrete cross-cutting and vertical guidance for AI in government. For instance, an inquiry by the Australian Parliamentary Joint Committee of Public Accounts and Audit (2025[28]) — which does not necessarily reflect the views of the Australian government — found that while some government entities are starting to adopt AI, they lack guidance in doing so. It recommended that a whole-of-government working group be established to consider what rules and governance frameworks necessary for AI systems across the public administration. However, there are notable exceptions, such as the "AI Playbook for the UK Government" (Chapter 4, Box 4.2) and France's structured approach to integrating AI in HRM (Box 5.19). Governments need to move beyond statements of intent to providing clear, practical guidance on AI adoption, investment, data governance, procurement and workforce development (Morley et al., 2019[29]).

This gap is cited as a challenge in many functions of government in Chapter 5. Overall, a lack of guidance can contribute to risk aversion, as guidance can clarify uncertainty and reduce doubts among civil servants. The need for guidance has been raised specifically in the functions of regulatory design, public procurement, fighting corruption and promoting public integrity, and tax administration to address legal and governance uncertainties that leave public servants uncertain if and how they can use AI in these functions. In some instances, guidance is needed to clearly interpret laws and regulations and their practical application. In others, they provide clarity in situations where formal rules may not yet exist. Given sectoral variations, cross-cutting AI approaches need to be complemented by specialised guidance accounting for unique policy challenges, risk profiles and data landscapes. Governments will also need to keep in mind that AI is a rapidly evolving field, and such guidance should be flexible and will likely need adaptation and iteration to keep up with the pace of change.

Further discussion on government actions to overcome this challenge can be found in Chapter 4, "Establishing key governance mechanisms and processes" and "Using policy levers to guide trustworthy AI".

Driving innovation while mitigating risks

OECD work last year ([30]) found that when it comes to AI in government, governments tend to focus on the negatives (i.e. AI risks and how to mitigate them), but not so much on the positives (i.e. AI benefits and how to take advantage of them). This is not unique to governments' approaches to internal AI activities. With regard to the broader economy and society, AI experts have found that government policy discussions and initiatives often recognise that AI may yield significant benefits, but government actions often do not explicitly target achievement of benefits. Rather, they indirectly address them through positive spillover effects when seeking to mitigate risks (OECD, 2024[31]). These experts have urged governments to take more direct action to seize the opportunities presented by AI.

Much of this risk-oriented focus can be attributed to risk aversion, which has long hindered digital government and public sector innovation efforts, fostering a culture resistant to change in which failure should be avoided at all costs, including with regard to AI (OECD, 2021[32]; 2017[3]; 2019[33]; Desouza, 2018[34]; SAS, 2020[35]; Richter, 2024[36]). A survey from Deloitte (2024[7]) showed that 63% of public sector leaders believed GenAI would erode the overall level of trust in national and global institutions. This caution contributed to slower adoption of AI in government than in industry. Instances of risk aversion can be seen in several government functions in Chapter 5, especially in public procurement, fighting corruption and promoting public integrity. For instance, integrity institutions can be risk averse due to fear of making mistakes in the AI adoption process — with government guidance emphasising *what not to do* rather than the provision of actionable guidance on *how to* adopt AI in a trustworthy manner. Issues also arise as a result of over-correction in response to AI incidents, especially ones covered in media. Examples include instances when government chatbots provided misinformation or were hacked (Hodges, 2024[37]; Fagan, 2024[38]). Risk aversion is also a common topic of conversation when discussing AI in government in meetings of relevant OECD working parties and fora.

Governments should pursue one of the generally accepted best practices to AI development in use, such as considering the level of potential risk or impact that an AI system might have and developing tailored and commensurate measures to overcome potential adverse issues. Yet governments often seem to treat most AI efforts as if they were of a high level of risk or impact, requiring exacting requirements across the board and imposing cumbersome bureaucratic requirements, daunting to public servants seeking to innovate. This can serve as sufficient disincentive to prevent the process of exploration and can be seen in a study from Deloitte (2023[39]), which analysed all policy initiatives in the OECD.AI Database of National Policies & Strategies.[3] It found risk-weighted policies, which aim to shift from one-size-fits-all approaches to a data-driven and risk-based approach, were rare, representing 2% of initiatives in the database. The problem is also recognised in the April 2025 policy from the US on "Accelerating Federal Use of AI through Innovation, Governance, and Public Trust" ([40]), which charges government agencies to "remove unnecessary and bureaucratic requirements that inhibit innovation and responsible adoption".

Public servants can also develop "algorithmic aversion". This is somewhat the opposite of "automation bias" discussed in Chapter 1, where "humans are reluctant to use algorithms despite their superior performance" (Cheng and Chouldechova, 2023[41]; Sunstein and Gaffe, 2024[42]), often after seeing mistakes in AI outputs. This suggests potential skills issues pertaining to the understanding of AI, its relative strengths and weaknesses and how to optimally use its outputs. It also suggests a lack of confidence in their abilities for human-machine collaboration as well as a lack of controlled environments for testing and safely experimenting with AI. These biases, which can distort perceptions of AI's reliability, can be mitigated through structured interventions, such as trainings on AI's strengths and limitations, as discussed further in Chapter 4 (Featherson, Shlonsky and Lewis, 2019[43]). Workers also need to feel they have a voice in inputs used for an AI system, and be able to use their professional judgement in how the outputs are used (Dietvorst, Simmons and Massey, 2018[44]; Cheng and Chouldechova, 2023[41]).

There is some evidence that risk aversion for AI in government may be receding as governments become more familiar with the technology. A recent study by Google Public Sector found that AI concerns among public IT leaders in the US over issues such as privacy and security are receding (Teale, 2025[45]). However, governments will need to be more active to overcome this risk-oriented focus to better consider trade-offs that better target opportunities. A variety of AI enablers and safeguards discussed in Chapter 4 can help divert from a culture of risk aversion and towards more controlled adoption and informed risk management.

Demonstrating results and return on investment

Governments have made significant strides in implementing AI solutions across various public domains, demonstrating tangible benefits in efficiency, accuracy and service delivery. However, monitoring of progress and thorough retrospective evaluation of impact remain underdeveloped aspects of government AI implementation. While isolated cases of success are well-documented, as seen below, comprehensive efforts to assess AI's contribution to public value creation are often lacking. This can be seen in the case of the UK, where "only 8% of AI projects show measurable benefits and only 16% show forecast costs, making it difficult to assess these against a cost-benefits analysis" (UK DSIT, 2025[21]). Specific to generative AI, a survey from Deloitte (2024[7]) in 14 countries shows that, despite anticipating increasing AI investments, 78% of government leaders surveyed report struggling to measure impacts from GenAI— significantly higher than those in other sectors, which poses a barrier to AI adoption and scaling even when other challenges, such as talent gaps, are resolved.

A handful of AI solutions have demonstrated concrete, measurable results that illustrate the technology's potential to transform government delivery. These quantifiable results provide valuable benchmarks for understanding AI's direct impact on operational efficiency and service quality:

- Peru's Amauta Pro AI system has transformed the speed at which courts can respond to victims of domestic violence. This AI-powered system has reduced the time needed to draft resolutions for protection measures from a lengthy 3 hours to 40 seconds. (see Box 5.63).
- In the EU, DATACROS Project developed a tool to detect anomalies in corporate ownership structures that may indicate risks of corruption, money laundering and other financial crimes. In 2021, the predictive tool correctly identified 83% of companies targeted by sanctions and 88% of companies with sanctioned owners (see Box 5.27).
- The US Federal Emergency Management Agency (FEMA) developed an AI system to assess structural damage across areas affected by Hurricane Ian, which reduced the number of structures requiring human review from over 1 million to just 77 000. Within 72 hours of the hurricane's landfall in 2022, FEMA had insights into the extent of damage across affected regions, enabling faster resource allocation and recovery planning. (see Box 5.58).

Particularly notable are cases where AI systems have been explicitly compared against human performance, highlighting significant improvements in speed, scale and resource utilisation that exceed human capability. In Singapore, government agencies transformed hiring with AI tools available on the market, enabling one agency to process over 3 000 applications for its Management Associate Program efficiently, saving EUR 44 000 (equivalent) and over 150 days of staff productivity.[4] Comparisons with human performance are important because they focus on the key counterfactual that is needed for evidence-based decision-making. Further, they push for a deeper understanding of human performance, making it possible to unveil implicit assumptions and biases that affect human delivery.

Beyond individual use cases, some governments have begun documenting AI's impact at organisational and national levels, revealing substantial financial benefits and operational improvements. These broader assessments help establish the cumulative value of AI investments across government functions. The Australian Taxation Office, for instance, reported that their AI approach combining real-time analytics, pre-filled forms, and anomaly detection systems helped protect approximately AUD 78.9 million in revenue across over 636 000 interactions with users in 2023-2024 (Box 5.5). Similarly, substantial results were observed in Austria with the activity of the Federal Ministry of Finance's Predictive Analytics Competence Centre (PACC), which made it possible to analyse 6.5 million cases across income, corporate and value-added tax sectors as well as customs transactions in 2023 (Box 5.3). These analyses detected instances of false reporting in employee tax assessments and identified fraudulent activities, resulting in additional tax revenues of approximately EUR 185 million. Looking to the future, a recent study by The Alan Turing Institute (2024[46]) on UK public services found AI could help automate 84% of the central government's service-related transactions, saving an equivalent of approximately 1 200 person-years of work every year.

Despite these successes, such considerations are rare, and governments face significant challenges in systematically monitoring AI's progress and impact. One key barrier is the lack of well-defined measurement and evaluation frameworks that can assess AI's contributions in a standardised manner. Many AI applications are integrated into complex administrative processes, making it difficult to isolate and measure their specific effects. Additionally, the challenge of benchmarking AI against human performance is compounded by the fact that many AI-enabled tasks would be infeasible or prohibitively time-consuming absent automation. There is also a limited understanding of the long-term impact of use of LLM's on human cognition, and whether their consistent use impacts the creativity, critical thinking skills and productivity of those that use them.

A final consideration is that different contexts may call for different methodologies. For instance, a theme of discussion at the latest OECD (2024[4]) Roundtable on Smart Cities was that cities need to explore different methodologies for measuring and evaluating success that align with their own objectives to allow them to set measurable goals. Some initial governments efforts to address these problems are emerging, such as the UK government publication on best practice for evaluating the impact of AI evaluation methods (Frontier economics, 2024[47]). The US (2025[48]) has also recently issued AI acquisitions policy that

recognises government agencies need to be "safeguarding taxpayer dollars by tracking AI performance and managing risks". Without robust monitoring mechanisms, governments risk wrongly estimating AI's value, potential risks and missing opportunities for improvement.

Establishing effective impact measurement frameworks is crucial for ensuring AI investments deliver real value to public administrations and citizens. As governments allocate increasing resources to AI development and deployment, demonstrating a clear return on investment (ROI) will become imperative. Reliable retrospective impact assessment mechanisms can help policymakers make informed decisions about scaling AI solutions, optimising resource allocation and justifying further funding. Furthermore, impact assessment provides essential feedback for refining AI systems and approaches, enabling continuous improvement cycles. Documented outputs also facilitate knowledge sharing across government entities, helping to scale successful approaches and avoid repeating unsuccessful ones. Perhaps most importantly, transparent reporting on AI's impacts — both positive and negative — is essential for maintaining public trust and accountability as these technologies become more deeply embedded in core government functions and activities. Different evaluation methods are appropriate for each context, but governments should try and compare the implementation of AI to the situation of its absence. The OECD has produced guidance on choosing an evaluation approach based on a variety of key considerations (Varazzani et al., 2023[49]).

Challenges that are somewhat less common or vary among government functions

Inflexible or outdated legal and regulatory environments

Inflexible, outdated or otherwise inadequate (e.g. excessive, lacking) regulatory environments pose many challenges. Many functions face regulatory or legal restrictions in data access and sharing, as discussed above. Beyond this, there can be confusion about AI accuracy and whether inadvertent errors introduced through the use of AI could lead to non-compliance with regulations and other rules, such as in fiscal reporting. Complexity in regulations is also a factor. For instance, tax administration officials face highly complex laws around tax processes, contributing to them largely relying on classic rules-based approaches. These challenges are as common at the local level as they are in national governments (OECD, 2024[4]).

Sometimes, issues with existing regulation is not the challenge, but gaps in regulation that lead to confusion over what is acceptable with AI. This confusion can contribute to other challenges, such as risk aversion or the preference for the maintenance of one's existing state of affairs (Samuleson and Zeckhauser, 1988[50]). For instance, because it is not specifically addressed in many countries, public procurement officials are often unclear on whether AI can be used in procurement processes, fearing that doing so could expose them to challenges from unsuccessful bidders or others who question the fairness of the process. This provides a general lack of incentive for change. Confusion also exists around whether using advanced AI systems, which are often highly capable but function in an opaque manner, can meet regulatory standards, such as International Standards on Auditing or evidentiary rules in the criminal justice system. In contrast, people may continue to operate without AI to forego these risks and also the benefits of AI use.

Regulatory environments pose a unique challenge regarding regulatory design and delivery. Beyond rules that restrict AI use, regulators should also be cautious and avoid making frequent changes to regulations, and to how they are implemented and enforced. Regulated entities need a level of clarity and predictability so they comply with regulations in a manner that causes minimal disruption with business operations. Frequent regulatory shifts, even if based on quality AI-informed insights, can lead to a volatile regulatory environment, making it difficult for businesses to plan long-term strategies and for the public to stay informed about current laws.

Governments can overcome these challenges by ensuring regulations and other formal rules are up to date, agile and minimise ambiguity. There is some evidence that governments are moving in this direction, with Deloitte (2023[39]) finding adaptive regulation — shifting from a "regulate and forget" mode to a responsive and interactive approach — to be one of the most common type of policies tracked by the OECD.AI Policy Observatory. Further guidance that helps contextualise regulations in particular functions of government is also important. Discussion on how governments are doing this can be found in Chapter 4, "Establishing key governance mechanisms and processes" and "Using policy levers to guide trustworthy AI".

High or uncertain costs of AI adoption and scaling

While AI adoption has the potential to reduce costs through enhanced productivity and efficiency, many government organisations struggle to make the initial financial investments to begin their AI adoption journeys, or to scale up use cases that prove successful. These costs can sometimes range from paying licensing fees per employee for service-based AI offerings, such as ChatGPT or Microsoft Copilot, to extensive development, customisation and support costs for more tailored or in-house solutions (Shark, 2025[51]; Barrett and Greene, 2024[52]). In the UK, a survey of government officials by SAS (2025[53]) found cost and budget restrictions to be the main challenge (raised by 67% of respondents), closely followed by a lack of internal skills (63%). Despite the critical nature of funding for AI, the OECD (2024[24]) Digital Government Index (DGI) highlights that only 15% of OECD countries have an investment framework in place for public AI investments.

Chapter 5 cites financial challenges in adopting AI for several functions of government, including in regulatory design and delivery, public services, tax administration, fighting corruption and promoting public integrity, and civic participation. In some instances, financial challenges also relate to the costs of recruiting or procuring skilled talent, with skills gaps discussed as a separate challenge above. Functions like tax administration have also indicated the process for securing budgets in government is a challenge.

This challenge can contribute to the OECD finding that governments often seem stuck in exploratory and pilot phases, with limited scaling of successful solutions. For instance, tax authorities have told the OECD that conducting small pilots is inexpensive and easy, even with advanced systems obtained from the private sector. However, such costs can grow exponentially as AI offerings are implemented more broadly within organisations or scaled up across other parts of government. Costs are particularly high for purpose-built solutions, with a group of 10 countries focusing on AI in the Public Interest stating that "the barrier to scaling AI models has been assumed to be primarily the lack of availability and affordability of compute" (France Élysée, 2025[54]).

Governments need to recognise that underinvestment in technology increases long-term costs and total costs of ownership (UK DSIT, 2025[21]). Some governments are seeking to address these issues through targeted investments, as well as the provisioning of central services that help to ease the need for each agency to build or buy their own solutions. These can be seen in Chapter 4 under "Investing purposefully", "Building out digital infrastructure" and "Creating spaces to experiment". Some are also using open-source models or exploring smaller models that can be designed to respond to specific societal and community needs, requiring less computational power and data (France Élysée, 2025[54]).

Although the costs governments and Chapter 5 highlight tend to focus on are financial, it is useful for governments to keep in mind that not only monetary costs impact AI adoption and scaling. Psychological costs related to AI use can also impact the extent to which individuals use AI tools in their day-to-day work, even if the investments are made to make them available. These costs can include search costs — which occur when people are searching for information but encounter outdated information, unclear language or confusing requirements — or cognitive costs, the mental resources people expend understanding complex information (Shahab and Lades, 2021[55]). "Sludge audits" are structured behavioural assessments of a decision-making process which aim to identify, prevent and reduce unnecessary frictions and

psychological costs, which prevent people from taking actions they would otherwise take (OECD, 2024[56]). By conducting sludge audits on the use of AI tools, governments can understand and address the barriers to acceptance that may limit AI adoption and scaling.

Governments can lack clarity on how much AI development and use can or should cost

The challenges discussed above are most relevant when such costs are known. Governments have reported to the OECD that there is often uncertainty or confusion about how much the development or use of different types of AI systems could or should cost. This makes it difficult for public institutions to plan effectively and evaluate vendor offerings when considering public procurement to source solutions. Gaining clarity on these costs can help ensure governments are prepared to adopt AI systems in a strategic and sustainable way. Yet, the OECD could identify no research discussing cost for the development or use of different types of AI systems in government. This suggests an optimal area for further research and analysis. Still, understanding costs in a broader sense or from specific government AI projects can help governments arrive at estimates for planning purposes. This section aims to take the first steps in helping governments do this, with the potential for more in-depth OECD work on this topic in the future.

The cost of adopting AI can vary significantly depending on the type of system and its scale of use. For instance, governments can pursue a variety of options in adopting AI, such as those touched on below.

Licensing private sector tools with fixed pricing per user or license

Companies offering AI tools often charge per user. This is the case for services such as Microsoft 365 Copilot, OpenAI's ChatGPT or Anthropic's Claude. Licenses for these services can be purchased at the enterprise level, with prices ranging between USD 30-100 per user per month. The enterprise version Microsoft 365 Copilot costs USD 30 per month, for example.[5] Thus, a government running a pilot project comparable to Australia's six-month pilot of 7 600 staff across 60 agencies (2024[57]) could estimate costs to be around USD 1.37 million for licenses alone; that does not include additions such as staff time for pilot administration and reporting results, as well as other overhead. For Australia's pilot, the Australian Treasury (2025[58]) estimated the license fee could pay for itself for a mid-level government staffer if it mitigated 13 minutes of their time per week for higher-value tasks. ChatGPT Enterprise costs are not published on OpenAI's website, although when asked, ChatGPT suggested a cost of USD 60-100 per user per month, varying based on volume and features. Anthropic's Claude Enterprise Plan pricing also depends on business needs and characteristics. While the prices are not available on Anthropic's website, third-party websites estimate a cost of USD 60 per person per month.[6]

Using private sector generative AI systems with volume-based pricing, such as tokens

GenAI systems are often offered through volume-based pricing, where the volume is calculated through API traffic. Whereas provision through licenses is more relevant when public servants use the AI tools directly (e.g. using Copilot to help draft documents), volume-based pricing is more relevant when governments build internal or public-facing services that interface with a proprietary model, or when they want to customise (i.e. fine-tune) the content the model considers or how it produces outputs. The main functions of the price include:

- Input tokens: tokens included in a prompt, such as instructions, context or data sent to the model.
- Output tokens: tokens generated by the model in response to an input.
- Training tokens: data (e.g. chunks of text) that an AI model learns from during training.

Models see data as tokens, not sentences or paragraphs. The cost of tokens depends on the company and may be based on the level of complexity and resources needed for an individual model. For instance, the costs for some commonly use models are in Table 3.1.

Table 3.1. Costs for 1 million tokens on common Generative AI models (USD)

One million tokens represent approximately 750 000 words, 100 000 lines of code, 11 hours of transcribed audio speech, or 1 hour of transcribed video

Model	Input tokens	Output	Training (if applicable)
OpenAI GPT-4o	USD 2.50	USD 10	
OpenAI GPT-4o (if fine-tuning to customise)	USD 3.75	USD 15	USD 25
OpenAI GPT-3.5-turbo	USD 0.50	USD 1.50	USD 8
Google Gemini 2.5 Pro	USD 1.25	USD 10	
Google Gemini 2.0 Flash	USD 0.10	USD 0.40	
Mistal Large 24.11	USD 2	USD 6	USD 9
Mistral NeMo	USD 0.15	USD 0.15	USD 1

Note: As of 10 April 2025. The inference costs for using a particular model or its equivalent tends to decrease over time (Stanford HAI, 2025[59]). Source: https://openai.com/api/pricing, https://ai.google.dev/gemini-api/docs/pricing, https://mistral.ai/products/la-plateforme, https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them, https://prompt.16x.engineer/blog/code-to-tokens-conversion, https://prompt.16x.engineer/blog/code-to-tokens-conversion, https://blog.google/technology/ai/google-gemini-next-generation-model-february-2024.

Further, some companies are offering foundation models through specific tiers dedicated to government agencies.[7] These services seek to meet governments' stringent security standards. Further, they aim to be tailored to the needs of governments, providing solutions that make it easier to manage their own security, privacy and compliance requirements, as well as enable them to use the services for activities that may fall outside the standard usage policies.

The experiences of one central government AI lab in pursuing this approach is discussed in Box 3.1. Proprietary models can also be used in concert with open-source models, as touched on below.

Box 3.1. Government AI lab's experience using proprietary AI

Operations and expenses

A central government AI lab in one county follows a phased approach to exploring, piloting and scaling AI projects for use in the public sector by civil servants. Overall, there may be as many as 100 projects being considered, around 15 undergoing limited testing, and around five to seven accessible to real users for a pilot or full deployment.

The lab uses cloud hosting and AI from Azure OpenAI, Vertex AI (Google), and Amazon Web Services (AWS). It has a budget of around EUR 17.5 million. All work is conducted in-house. Most of its expenses are staff costs, including roughly:

- 15 full-time equivalent staff (FTE) for technical talent (seven for development and AI engineering, four for design and user research, four for cloud/infrastructure).
- Six FTEs for delivery management, which is critical for ensuring technical talent can focus on technical challenges, while delivery managers focus on addressing challenges regarding policy and bureaucracy.
- Six FTEs for impact analysts who use data science to study project results and impact.

Its largest deployed project has around 4 000 users, with around eight FTEs working on it. Other projects are smaller, with some having one or two FTEs. Overall, the lab's products have around 10 000 monthly users. The total costs for AI cloud services, including tokens, are around EUR 3 500 per month.

> **Lessons learned**
>
> - The first few projects are by far the most expensive and time consuming, with significant investments in setting up cloud infrastructure and deployment templates that can be easily re-used for future projects. Deployments that took three weeks each for the first few projects now take 30 minutes.
>
> - Having robust cloud infrastructure in place is important, optimally shared among projects to promote synergies.
>
> - The lab considered the pros and cons of using proprietary AI models versus custom deployments of open-source models (e.g. Meta's Llama). It determined volume-priced proprietary models to be more effective because civil servants tend to use the AI systems from 9:00-18:00. For a custom deployment, they would need to pay for GPU usage all day, even when the models are not being used. Overall, token-based pricing was less expensive for their needs. In addition, this approach allows the lab to spin-up new model instances more quickly and easily. For instance, it can deploy a new GPT-4.1 model in around five minutes, whereas custom deploying an open-source model could take weeks of infrastructure work.
>   - Overall, the lab estimated it would have cost EUR 9 300 per month to self-host host a Llama model, where they current spend around EUR 3 500 for tokens.
>
> - As usage of the lab's AI tools has grown, the lab is reaching the limits of what cloud providers are willing to provide in terms of pay-as-you-go pricing. As their usage continues to increase, they face the choice of either 1) purchasing GPU capacity from the cloud providers, or 2) self-hosting open-source models and paying for GPU access directly (as mentioned in the previous bullet). For the lab, option 1 may be optimal because it still allows for rapid deployment.
>
> - The cost of developing AI, in terms of both technical resources (e.g. cloud services, tokens) and human resources is decreasing rapidly. The lab is finding that they can increasingly use AI to build AI, potentially changing labour demands. The implications for this have yet to be determined.
>
> Source: OECD interview with officials from an undisclosed country on 18 April 2025. The OECD is not publishing the name of the country or lab because of the preliminary nature of the estimates and analysis.

Developing narrow, purpose-built custom ML applications (either in-house or procured)

Narrow AI systems tailored for specific public-sector tasks can range from relatively small expenses to multi-million-dollar projects**.** These systems involve ML approaches developed for a specific use case, such as fraud detection, traffic optimisation or document classification. Simple pilots might be built for a few thousand dollars, whereas complex national systems can cost millions or tens of millions (USD), especially if scoping in defence applications (Barnett, 2020[60]). As one example, the South Australian government is piloting four AI-enabled cameras aimed at cutting traffic congestion by analysing congestion and adjusting traffic light cycles at a cost of USD 218 000 (equivalent) (Jackson, 2025[61]).

However, costs vary widely depending on complexity and context, and additional costs may be needed for data preparation, infrastructure and ongoing monitoring and maintenance. Because approaches and associated costs vary significant depending on the use case, it is difficult to provide estimates beyond these examples. Further analysis may be warranted to consider different aspects of such use cases and what different governments around the word have paid for them.

Developing systems using pre-trained open-source models

Compared to the custom development and training of a GenAI model, discussed below, pre-trained open-source AI models (such as Meta's Llama models) can offer reduced costs with government still able to highly customise the model to meet their needs.[8] Open-source models can be self-hosted either on the cloud or on-premises, offering governments greater control over their data and long-term cost efficiency. While leveraging a pre-trained open-source model can reduce training costs to those incurred for fine-tuning, still other remaining costs can be significant. For instance, self-hosting eliminates recurrent fees for tokens or licenses, but it requires significant upfront investment in hardware and infrastructure, cloud resources, energy consumption, and maintenance and support costs.

Despite the higher initial investment compared to licence or volume-based pricing, some governments have found that self-hosting can be cost-effective at scale and unlock use cases not feasible to them through commercial APIs (e.g. sensitive intelligence tasks, always-on local services, or offline operations in critical infrastructure). For instance, Chinese Taipei invested USD 7.4 million to develop its own foundation model called Trustworthy AI Dialogue Engine (TAIDE), which uses Meta's Llama open-source models (Creery, 2024[62]).[9]

Government use of the open-source platform Polis (Box 5.36) represents a narrower use of open-source AI than the TAIDE efforts to develop a foundation model. One government organisation that built a customised and self-deployed version of Polis for a large-scale public engagement campaign, consisting of 33 regional and national Polis discussions with 30 000 participations, incurred overall costs of around EUR 422 500 over a 14-month period.[10] These expenses consisted of EUR 195 000 for expert web development, cloud services and an outsourced user experience (UX) co-design sprint, and EUR 227 500 in staff time to implement Polis into the organisation's existing workflows and to coordinate and conduct civic engagement activities. The same organisation has since invested EUR 200 000 in further enhancing their Polis application, split 50/50 between technical and staff costs. This enhancement work included user interface (UI) design and implementation and other technical development work, which have been open sourced for other Polis users. Overall, the organisation estimates that two full-time equivalent staff (FTE) is sufficient to manage the work, with a mix of expertise needed (such as technical development, project management, UX design and digital participation).

Open-source models can also be used in conjunction with proprietary models. For instance, one of the virtual assistants discussed in Chapter 5 uses both Google's Gemini 1.5 Flash and a Llama model from Meta, with the chat interface and orchestration of the two models developed in-house using open-source technologies.[11] The system is being piloted with 18 000 users, with the main costs being associated with the use of the Gemini LLM and cloud web hosting for the chat application. While the LLM platform costs the government around EUR 18 000 per month, they expect a substantial reduction as they understand how to use the model most efficiently. The web hosting costs around EUR 2 300 per month. Overall, they estimate their costs at EUR 0.93-1.55 per user per month. Officials are in early discussions to scale the pilot up in other departments. The development and coordination team consists of approximately 10 FTEs.

Developing custom GenAI models built and trained from scratch

Developing custom built and trained GenAI models is generally the most expensive option (for comparable performance) due to high initial investment and operational complexity. Costs of training an LLM depend on model size (larger models with more parameters require more computational power and consume more energy), the quality and quantity of the training data (influencing the cost of data acquisition and curation), choices of infrastructure (whether training occurs on premises or cloud based), and the efficiency of training algorithms used.

AI companies often do not publicly disclose the training costs associated with their models, although researchers estimate that current popular models cost around USD 41-192 million (Stanford HAI, 2025[59]).

With the cost of training state-of-the-art models increasing two- to threefold each year, some research estimates that training the largest models may cost over EUR 1 billion by 2027 (Cottier et al., 2024[63]). While these costs can seem high, they can pale in comparison to the significant research and development investments, staff costs and data gathering efforts needed to achieve the latest foundation models (Stanford HAI, 2025[59]). Leading AI companies also have other prerequisites for developing such models, such as deep technical talent, and often, strategic partnerships with other companies.

Yet governments do not necessarily need to embark on building such extensive and powerful systems that seek to beat market competitors. Training from scratch of government-funded LLMs, for instance, can require fewer staff and cost resources, especially for those with fewer parameters or seeking to maximise relevance for a specific country, region or language. For instance, OpenEuroLLM has a total budget of EUR 37.4 million, implying a fraction of this sum will be dedicated to training its foundation model with another fraction for staff (EC, 2025[64]).[12] In another example, one European country custom developed and trained an LLM in the national language, with the total cost, including personnel, coming in around EUR 500 000, of which EUR 300 000 was dedicated to GPU usage.[13] Another example is GPT-NL in the Netherlands, which is investing around EUR 13.5 million provided by the Netherlands Ministry of Economic Affairs and Climate (EZK) to train a model (2023[65]). Other efforts have been undertaken in in Japan, Singapore, Spain, Sweden and the United Arab Emirates (Chavez, 2024[66]). The collaboratively developed BigScience Large Open-science Open-access Multilingual Language Model (BLOOM)[14] involved significant contributions from government agencies. The project was primarily led by Hugging Face and the French National Centre for Scientific Research (CNRS), supported by a public compute grant on the French public supercomputer "Jean Zay". The estimated cost of training is USD 2-5 million.

Outdated legacy information technology systems

Many governments' AI ambitions are slow to materialise because of outdated legacy IT systems not suitable for AI development or use, or inadequate to manage and exchange large amounts of quality, interoperable data (Irani et al., 2023[67]). Such systems can result in significant missed opportunities. For example, in the UK alone, "taxpayer funded services from the NHS to local councils are missing out on GBP 45 billion in productivity savings — more than enough to pay for every primary school in the UK for a full year — because they are too often dependent on old and outdated technology" (UK DSIT, 2025[68]). The government (2025[21]) estimates 28% of central government IT systems are outdated, reaching 70% in some organisations, and 57% of UK government officials surveyed by software company SAS (2025[53]) cited legacy systems as a barrier to AI adoption. The issue has also been raised by the UK PAC (2025[6]) as an impediment to the use of AI in government.

Chapter 5 discusses how outdated legacy technology impacts AI adoption. For instance, the potential for AI in public financial management is limited by outdated financial management information systems in governments around the world, with such systems exceeding a decade old in most OECD countries (Rivero del Paso et al., 2023[69]; OECD, 2024[70]). Despite the significance of the challenge, considerations and analysis for legacy technology's adverse effects in AI adoption appear light in most countries and government functions. While the previous paragraph includes significant detail about the scale of the challenge in the UK to illustrate the point, this is largely because most other governments have not conducted the analysis necessary to articulate the problem in such a manner.

This challenge is dependent on others, including the significant costs of funding the remediation of legacy systems. Outdated legacy technology also contributes to other challenges, such as data issues and an "overreliance on contractors sending costs rocketing", including to maintain outdated systems, with "maintenance of legacy systems costing often three to four times that of modern alternatives" (UK DSIT, 2025[68]). These expenses could be better placed in innovation and modernisation efforts.

Governments are taking a variety of measures to modernise their systems to be more AI-ready. In a novel instance, the US Department of Defence is using AI to modernise legacy code (Harper, 2024[71]). More

traditionally, some governments are providing targeted funding for modernisation efforts (see Chapter 4, "Funding AI and supporting coherent investments across government").

To overcome the implementation challenges outlined in this chapter, as well as mitigate the risks outlined in Chapter 1, governments can turn to policy. Indeed, some governments are already doing so. The following chapter looks at policy measures governments can take — and actions some are already taking — to deliver AI that is trustworthy and to benefit from AI's full potential.

References

Austin, T. et al. (2024), *A snapshot of how public sector leaders feel about generative AI*, https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-adoption-in-public-sector.html. [7]

Australia DTA (2024), *Evaluation of whole-of-government trial into generative AI: Now available*, https://www.dta.gov.au/blogs/evaluation-whole-government-trial-generative-ai-now-available. [57]

Australia Treasury (2025), *Evaluation of a trial of generative AI (Copilot) in The Treasury*, https://evaluation.treasury.gov.au/publications/evaluation-generative-artificial-intelligence. [58]

Autio, C., K. Communigs and B. Elliott (2023), *A Snapshot of Artificial Intelligence Procurement Challenges*, https://files.thegovlab.org/a-snapshot-of-ai-procurement-challenges-june2023.pdf. [20]

Barnett, J. (2020), *JAIC awards biggest contract yet to buy AI for the battlefield*, https://fedscoop.com/battlefield-ai-jaic-booz-allen. [60]

Barrett, K. and R. Greene (2024), *The Future of AI For the Public Sector: The Challenges and Solutions*, https://www.businessofgovernment.org/blog/future-ai-public-sector-challenges-and-solutions. [52]

Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, European Commission, https://op.europa.eu/s/z4XY. [2]

Chavez, P. (2024), *Sovereign AI in a Hybrid World: National Strategies and Policy Responses*, https://www.lawfaremedia.org/article/sovereign-ai-in-a-hybrid-world--national-strategies-and-policy-responses. [66]

Cheng, L. and A. Chouldechova (2023), "Overcoming Algorithm Aversion: A Comparison between Process and Outcome Control", *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1-27, https://doi.org/10.1145/3544548.3581253. [41]

Cottier, B. et al. (2024), *The rising costs of training frontier AI models*, https://arxiv.org/abs/2405.21015. [63]

Creery, J. (2024), *Taiwan Builds Own AI Language Model to Counter China's Influence*, https://www.bloomberg.com/news/articles/2024-01-25/taiwan-builds-own-ai-language-model-to-counter-china-s-influence?. [62]

de Mello, L. and T. Ter-Minassian (2020), "Digitalisation challenges and opportunities for subnational governments"*, OECD Working Papers on Fiscal Federalism*, No. 31, OECD Publishing, Paris, https://doi.org/10.1787/9582594a-en. [17]

Desouza, K. (2018), *Delivering Artificial Intelligence in Government: Challenges and opportunities*, https://www.businessofgovernment.org/sites/default/files/Delivering%20Artificial%20Intelligence%20in%20Government.pdf. [34]

Dietvorst, B., J. Simmons and C. Massey (2018), "Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them", *Management Science*, Vol. 64/3, pp. 1155-1170, https://doi.org/10.1287/mnsc.2016.2643. [44]

EC (2025), *A pioneering AI project awarded for opening Large Language Models to European languages*, https://digital-strategy.ec.europa.eu/en/news/pioneering-ai-project-awarded-opening-large-language-models-european-languages. [64]

EC (2024), *Adoption of AI, blockchain and other emerging technologies within the European public sector – A public sector Tech Watch report*, Publications Office of the European Union, https://data.europa.eu/doi/10.2799/3438251. [8]

EC (2024), *Public Sector Tech Watch latest dataset of selected cases*, http://data.europa.eu/89h/e8e7bddd-8510-4936-9fa6-7e1b399cbd92 (accessed on 4 March 2025). [9]

Fagan, M. (2024), *AI for the People: Use Cases for Government*, https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/working.papers/M-RCBG%20Working%20Paper%202024-02_AI%20for%20the%20People.pdf. [38]

Featherson, R., A. Shlonsky and C. Lewis (2019), "Intervetions to Mitigate Bias in Social Work Decision-Making: A Sytematic Review", *Research on Social Work Practice*, Vol. 29/7, https://doi.org/10.1177/1049731518819160. [43]

France Élysée (2025), *The Paris Charter on Artificial Intelligence in the Public Interest*, https://www.elysee.fr/en/emmanuel-macron/2025/02/11/the-paris-charter-on-artificial-intelligence-in-the-public-interest. [54]

Frontier economics (2024), *Guidance on the impact of AI interventions*, https://www.frontier-economics.com/uk/en/news-and-insights/news/news-article-i21121-analysing-the-impact-of-ai-interventions-in-government. [47]

Government of the Netherlands (2023), *The Netherlands is building its own open language model GPT-NL*, https://www.digitaleoverheid.nl/nieuws/nederland-bouwt-eigen-open-taalmodel-gpt-nl/. [65]

Harper, J. (2024), *Pentagon using AI to modernize legacy code*, https://defensescoop.com/2024/09/12/pentagon-artificial-intelligence-modernize-legacy-code-john-hale/. [71]

Hodges, D. (2024), *Fumbles can't kill the government's AI appetite*, https://www.themandarin.com.au/249756-red-faces-and-fumbles-cant-kill-governments-ai-appetite/. [37]

Irani, Z. et al. (2023), "The impact of legacy systems on digital transformation in European public administration: Lesson learned from a multi case analysis", *Government Information Quarterly*, Vol. 40/1, p. 101784, https://doi.org/10.1016/j.giq.2022.101784. [67]

Jackson, B. (2025), *South Australian drivers to be monitored by AI cameras*, https://www.news.com.au/technology/south-australian-drivers-to-be-monitored-by-ai-cameras/news-story/8a63be3e80d4bba60735d58ec8c473db?utm_source=chatgpt.com. [61]

Mariani, J., W. Eggers and P. Kishnani (2023), *The AI regulations that aren't being talked about*, https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-regulations-around-the-world.html. [39]

Mitchell, S. (2025), *Skills gap in public sector IT fuels outsourcing reliance*, https://itbrief.co.uk/story/skills-gap-in-public-sector-it-fuels-outsourcing-reliance. [19]

Monteiro, B., A. Hlacs and P. Boéchat (2024), "Public procurement for public sector innovation: Facilitating innovators' access to innovation procurement", *OECD Working Papers on Public Governance*, No. 80, OECD Publishing, Paris, https://doi.org/10.1787/9aad76b7-en. [27]

Morley, J. et al. (2019), "From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices", *Science and Engineering Ethics*, Vol. 26/4, pp. 2141-2168, https://doi.org/10.1007/s11948-019-00165-5. [29]

Muñoz-Cadena, S. et al. (2025), *Sistemas de IA en el sector público de América Latina y el Caribe (Versión V2)*, https://sistemaspublicos.tech/sistemas-de-ia-en-america-latina/ (accessed on 29 April 2025). [10]

NASCIO (2024), *Generative Artificial Intelligence and its Impact on State Government IT Workforces*, National Association of State Chief Information Officers, https://www.nascio.org/resource-center/resources/generative-artificial-intelligence-and-its-impact-on-state-government-it-workforces/. [15]

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. [24]

OECD (2024), "Assessing potential future artificial intelligence risks, benefits and policy imperatives", *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, https://doi.org/10.1787/3f4e3dfb-en. [31]

OECD (2024), "Financial Management Information Systems in OECD countries", *OECD Papers on Budgeting*, No. 2024/02, OECD Publishing, Paris, https://doi.org/10.1787/ce8367cd-en. [70]

OECD (2024), *Fixing Frictions: 'Sludge audits' around the world*, OECD Publihing, https://doi.org/10.1787/14e1c5e8-en-fr. [56]

OECD (2024), "Governing with Artificial Intelligence: Are governments ready?", *OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, https://doi.org/10.1787/26324bc2-en. [30]

OECD (2024), *Shaping smart cities of all sizes*, OECD Publishing, https://www.oecd.org/content/dam/oecd/en/about/programmes/cfe/the-oecd-programme-on-smart-cities-and-inclusive-growth/Proceedings-4th-Roundtable-Smart-Cities-Inclusive-Growth.pdf/_jcr_content/renditions/original./Proceedings-4th-Roundtable-Smart-Cities-In. [4]

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. [23]

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. [32]

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. [25]

OECD (2017), *Fostering Innovation in the Public Sector*, OECD Publishing, Paris, https://doi.org/10.1787/9789264270879-en. [3]

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. [1]

OECD/UNESCO (2024), *G7 Toolkit for Artificial Intelligence in the Public Sector*, OECD Publishing, Paris, https://doi.org/10.1787/421c1244-en. [11]

Parliament of Australia (2025), *Report 510: Inquiry into the use and governance of artificial intelligence systems by public sector entities - 'Proceed with Caution'*, https://parlinfo.aph.gov.au/parlInfo/download/committees/reportjnt/RB000567/toc_pdf/Report510Inquiryintotheuseandgovernanceofartificialintelligencesystemsbypublicsectorentities-'ProceedwithCaution'.pdf. [28]

PTI (2024), *AI and City/County Government: Survey Results*, Public Technology Institute, https://fusionlp.org/wp-content/uploads/2024/11/AI-Survey-City-and-County-Final-2024.pdf. [16]

Richter, A. (2024), *Navigating Generative AI in Government*, https://businessofgovernment.org/report/navigating-generative-ai-government. [36]

Rivero del Paso, L. et al. (2023), *Digital Solutions Guidelines for Public Financial Management*, https://www.imf.org/en/Publications/TNM/Issues/2023/10/06/Digital-Solutions-Guidelines-for-Public-Financial-Management-537781. [69]

Ryseff, J., B. De Bruhl and S. Newberry (2024), *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*, RAND, https://www.rand.org/pubs/research_reports/RRA2680-1.html. [5]

Ryseff, J. and A. Narayanan (2025), *Why AI Projects Fail*, https://www.rand.org/pubs/presentations/PTA2680-1.html. [22]

Salesforce (2024), *6 in 10 IT Workers Report Shortage of AI Skills in the Public Sector*, https://www.salesforce.com/news/stories/public-sector-ai-statistics/. [12]

Samuleson, W. and W. Zeckhauser (1988), "Status quo bias in decision making", *Journal of Risk and Uncertainty*, Vol. 1, pp. 7-59, https://doi.org/10.1007/bf00055564. [50]

SAS (2025), *Slow uptake of AI in government hindering strategic goals, new research finds*, https://www.sas.com/en_gb/news/press-releases/2024/september/slow-uptake-of-ai-in-government-hindering-strategic-goals.html. [53]

SAS (2020), *AI in government: The path to adoption and deployment*, https://www.sas.com/en_sa/insights/articles/analytics/ai-in-government.html. [35]

Shahab, S. and L. Lades (2021), "Sludge and transaction costs", *Behaviorual Public Policy*, Vol. 1/22, https://doi.org/10.1017/bpp.2021.12. [55]

Shark, A. (2025), *What the Rising Costs of AI Means for Government*, https://statetechmagazine.com/article/2025/01/what-rising-costs-ai-means-government. [51]

Stanford HAI (2025), *Artificial Intelligence Index Report 2025*, https://hai-production.s3.amazonaws.com/files/hai_ai_index_report_2025.pdf. [59]

Sunstein, C. and J. Gaffe (2024), *An Anatomy of Algorithm Aversion*, Elsevier BV, https://doi.org/10.2139/ssrn.4865492. [42]

Teale, C. (2025), *Public-sector concerns over AI are lessening, survey says*, https://www.route-fifty.com/artificial-intelligence/2025/02/public-sector-concerns-over-ai-are-lessening-survey-says/403328. [45]

The Alan Turing Institute (2024), *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions*, https://www.turing.ac.uk/news/publications/ai-bureaucratic-productivity-measuring-potential-ai-help-automate-143-million-uk. [46]

Trajkovski, G. (2024), "Bridging the public administration-AI divide: A skills perspective", *Public Administration and Development*, Vol. 44/5, pp. 412-426, https://doi.org/10.1002/pad.2061. [18]

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. [33]

UK Committee of Public Accounts (2025), *Use of AI in Government*, https://committees.parliament.uk/publications/47199/documents/244683/default/. [6]

UK DSIT (2025), *Archaic tech sees public sector miss £45 billion annual savings*, https://www.gov.uk/government/news/archaic-tech-sees-public-sector-miss-45-billion-annual-savings. [68]

UK DSIT (2025), *State of digital government review*, https://www.gov.uk/government/publications/state-of-digital-government-review/state-of-digital-government-review. [21]

UK NAO (2024), *Use of artificial intelligence in government*, National Audit Office, https://www.nao.org.uk/wp-content/uploads/2024/03/use-of-artificial-intelligence-in-government.pdf. [13]

UN Habitat (2024), *Global assessment of Responsiuble AI in cities*, https://unhabitat.org/sites/default/files/2024/08/global_assessment_of_responsible_ai_in_cities_21082024.pdf. [14]

US OMB (2025), *Accelerating Federal Use of AI through Innovation, Governance, and Public Trust*, https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf. [40]

US OMB (2025), *Driving Efficient Acquisition of Artificial Intelligence in Government*, White House Office of Management and Bydget, https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf. [48]

van Noordt, C., R. Medaglia and L. Tangi (2023), "Policy initiatives for Artificial Intelligence-enabled government: An analysis of national strategies in Europe", *Public Policy and Administration*, https://doi.org/10.1177/09520767231198411. [26]

Varazzani, C. et al. (2023), "Seven routes to experimentation in policymaking: A guide to applied behavioural science methods", *OECD Working Papers on Public Governance*, Vol. 64, https://doi.org/10.1787/918b6a04-en. [49]

Waters, R. (2024), *Meta under fire for 'polluting' open-source*, https://www.ft.com/content/397c50d8-8796-4042-a814-0ac2c068361f. [72]

Notes

[1] "Salesforce conducted a double-anonymous survey of 600 IT professionals (200 IT leaders and 400 IT individual contributors) in Australia, France, Germany, the United Kingdom and the United States. Respondents work across industries, including technology, financial services, media and entertainment, manufacturing, retail, healthcare, the public sector and more. The survey was fielded in December 2023 and January 2024" (2024[12]).

[2] The source for this sentence is a report by the Joint Committee of Public Accounts and Audit, Parliament of Australia (2025[28]). The findings of the Committee and are not necessarily representative of the Australian Government's views

[3] https://oecd.ai/dashboards.

[4] https://impress.ai/case-studies/publicsector.

[5] Based on Microsoft's US-oriented website (https://www.microsoft.com/en-us/microsoft-365/copilot) as of 10 April 2024. Annual subscription pricing.

[6] https://team-gpt.com/blog/claude-pricing.

[7] See, for example, https://openai.com/global-affairs/introducing-chatgpt-gov and https://www.anthropic.com/news/expanding-access-to-claude-for-government.

[8] The use of "open-source" models for this report does not imply that such models are released under an open-source license approved by the Open Source Initiative (OSI), a nonprofit steward of The Open Source Definition (https://opensource.org/osd). OSI has criticised some companies that call their models open source because they only provide the weights for the model, and not other elements, such as the training

data, code and training practices (Waters, 2024[72]). Some argue that such models should be called "open weight" instead of "open source".

[9] https://en.taide.tw.

[10] Information provided by an undisclosed country to the OECD. The OECD is not publishing the name of the country or project because of the preliminary nature of the estimates and analysis.

[11] Information provided by an undisclosed country to the OECD. The OECD is not publishing the name of the country or project because of the preliminary nature of the estimates and analysis.

[12] https://openeurollm.eu.

[13] Figures reported to the OECD by a non-disclosed country.

[14] https://huggingface.co/bigscience/bloom.

4 Enablers, guardrails and engagement for unlocking trustworthy AI

This chapter sets out how governments can realise AI's potential while managing risks. It establishes three pillars — enablers, guardrails and engagement — that together form the OECD Framework for Trustworthy AI in Government. Enablers include governance, data, digital infrastructure, skills and talent, purposeful investment, procurement and partnerships; guardrails cover non-binding and binding instruments, transparency and risk management, and oversight; engagement spans citizens, civil servants and cross-border collaboration. The chapter calls for a systems approach, proportionate, risk-based application of measures, and practical mechanisms such as experimentation and impact assessment and auditing to support trustworthy adoption at scale.

Key messages

- Governance and policy initiatives can help governments to fully exploit AI's potential and address its various risks and implementation challenges. Governments should take a systems approach and seek to anticipate future changes. Proposed measures should strengthen enablers, establish guardrails and engage with stakeholders.
- Enablers include establishing key governance mechanisms and processes, understanding data's role as the foundation for AI, building digital infrastructure, fostering skills and talent, investing purposefully, effectively using public procurement and expanding AI's potential through partnerships.
- Guardrails can be binding and non-binding policy levers, transparency processes and accountability mechanisms.
- Engagement with stakeholders can take the form of citizen assemblies, engaging with civil servants, involving users in AI development and collaborating across borders.
- Taken together, these policy measures form a Framework for Trustworthy AI in Government, which can help governments to align their actions with the OECD AI Principles. Future OECD work will address elements of the framework more in-depth.

Policy action to unlock AI's potential

While the rest of this report discusses governments' opportunities and challenges in *governing with AI*, this chapter focuses on how to realise success through *governing AI* in government. To fully leverage the potential of AI in the government while mitigating risks, governments need to take an intentional, strategic and responsible approach. This approach should align with the OECD AI Principles, but contextually specific and appropriate for the development and use of AI for and by governments. In particular, governments should pursue three courses of action:

1. strengthening **enablers** (e.g. quality data, digital and AI skills, funding and digital infrastructure) to overcome key implementation challenges and deliver the expected results
2. establishing **guardrails** (e.g. transparency, accountability and risk management tools) to anticipate and manage associated risks, and
3. fostering **engagement** with stakeholders (including the public) to develop AI systems that take their needs into account.

These actions aim to harness opportunities and address the various risks and implementation challenges associated with AI through targeted policy measures. For instance, issues related to the need for sufficient and quality data are addressed through enablers that focus on building robust data governance and infrastructure. In another example, insufficient guidance and outdated regulations can be addressed through guardrails such as binding and non-binding policy levers, including agile regulatory instruments, helping to ensure AI operates within clearly established ethical, operational and legal boundaries. Active stakeholder engagement throughout the AI system lifecycle (development, deployment and use of AI technologies) and the policy cycle (for designing, implementing and evaluating AI governance and policy) can enrich the understanding, attitudes and behaviours of stakeholders, including the public, and align technology and governance developments with societal needs.

The sections below seek to address the main points of attention for governments to create an environment that enables the strategic and responsible use of AI across government systems and functions. They

provide a comprehensive analysis of specific policy actions and priorities along the three action areas. Each section outlines key policy options that governments should consider for a coherent and sustainable approach. Governments should consider these options in light of their own context, including their current digital government maturity. For instance, instead of seeking to put in place all of the items discussed at once, governments could consider progressive governance and AI adoption roadmaps based on institutional, cultural and technological capacities.

A holistic, systems approach can maximise the value of AI in government

In establishing the enablers, guardrails and engagement mechanisms discussed in this chapter, governments should take a systems approach to AI, seeking to understand and address public problems by viewing them as part of a larger, interconnected system rather than in isolation (OECD, 2017[1]). Traditionally, public policymakers have addressed social problems through discrete interventions layered on top of one another, building on a "cause and effect" relationship. However, these interventions may shift consequences from one part of the system to another, simply addressing symptoms while ignoring causes. AI represents an opportunity to reimagine how government works, in terms of both internal operations and public-facing services. Governments should think beyond how AI can fit within existing government systems and structures; they need to think about how AI can contribute to entirely rethinking processes and systems. Otherwise, governments run the risk of simply automating inefficiency and further reinforcing misaligned incentives and governance approaches. Emerging practices can assist, such as "sludge audits", which are structured behavioural assessments of a process to identify frictions that result in people being less likely to complete the process or expend undue psychological effort while doing so (OECD, 2024[2]). The OECD has a dedicated line of work to systems approaches that can further assist.[1]

Governments should recognise the potential for and seek to anticipate future changes

There is still much to learn about AI and much remains unknown about its ongoing evolution. Establishing the enablers, guardrails and engagement mechanisms discussed in this chapter may only take into account today's knowledge — what is known about the uses and implications of current AI and about the potential for tomorrow. Yet, there are major unknowns that will only be resolved over time as the technology develops and its potential uses explored. AI in government will be an ongoing journey of discovery, both welcome and unwelcome, and unexpected and unintended developments. Governments should employ an agile and adaptive approach to adjust to new opportunities and changing behaviours. Many tasks AI cannot satisfy today will likely become feasible in the future. AI strategies and frameworks should be flexible enough to evolve with changing capabilities and contexts. Governments need to improve their early engagement with weak signals that indicate how the future may transpire. This will enable them to understand where and when to best intervene, without waiting for processes and trends to become established, and thus expensive and difficult to shift.

These considerations are relevant not only for AI governance, but also for its use in government. The use cases discussed in this report (in-depth in Chapter 5 and synthesised in Chapter 2) generally represent incremental improvements and productivity gains. This, however, should not obscure that emergent or future uses of AI could be completely new or handle previously impossible, impractical or even inconceivable tasks and create new opportunities and risks for government. This can lead to use cases and governance approaches that will need to be created, reinvented or stopped. OECD efforts on Anticipatory Innovation Governance (AIG) and Strategic Foresight can help government better understand and shape potential AI futures.[2]

Strengthening enablers to facilitate the adoption of trustworthy AI

Enablers are the foundational elements and resources necessary for AI implementation in government.[3] They create an environment where skilled public servants can effectively and reliably design and deploy AI. Their practical support allows government institutions to fully harness AI's potential. The sections below review seven key enablers: **governance, data (including open government data), digital infrastructure, skills and talent, AI investments, public procurement and partnering with non-governmental actors**. These were initially defined by the OECD in 2024 ([3]) and are further developed in the sections below. Each section considers policy options governments can adopt to deploy these enablers in their contexts, drawing on international best practices.

Establishing key governance mechanisms and processes

Governments are accelerating their adoption of AI, and in most cases, have outlined their goals in national AI strategies. Yet, this adoption is often piecemeal in practice at a high-level, without establishing comprehensive and robust governance arrangements to ensure AI's responsible use, long-term impact and sustainability. As AI becomes increasingly integrated into government operations, robust governance arrangements should be established domestically to ensure the trustworthy, sustainable and effective use of AI. They can also promote a clear narrative of the benefits of AI to build support within and outside government.

Using strong leadership for a cohesive vision

Strong leadership is a critical factor in achieving AI adoption in government. It is vital to setting the right tone from the highest levels of government and actively communicating the potential benefits of AI (Berryhill et al., 2019[4]). While establishing strategies and principles to help ensure trustworthy AI adoption is critical, solid and effective leadership can build a cohesive vision for AI and set a "tone at the top" that builds confidence in AI, both within and beyond government. For instance, in the United Kingdom (UK), the past two prime ministers (PMs), Rishi Sunak and Keir Starmer, have championed AI's adoption both inside and outside government. Under Sunak's leadership, the UK catalysed global attention and international collaboration on AI through convening the first global AI Summit in November 2023, with subsequent events organised by Korea and France. Each has concluded with a declaration signed by many governments outlining AI-related commitments.[4] More targeted on government, the UK's Incubator for AI (i.AI), which aims to improve lives, drive growth and deliver better public services, was also launched during Sunak's tenure. Under current PM Starmer's leadership, the UK (2025[5]) launched the AI Playbook for the UK Government (Box 4.2) and has put forth a bold plan to use AI in "reshaping the state to make it work for working people", including through the creation of 2 000 tech and AI apprentices in government. Finally, in cascading strong leadership throughout UK government organisations, (2025[6]) "A blueprint for modern digital government" requires all public sector organisations to include a digital leader on their executive committee by 2026.

Those at the top have the power to set a strategic direction that can permeate levels below, helping to frame the use of AI within the culture at large. As stated in the OECD *Framework for Digital Talent and Skills in the Public Sector* (2021[7]), "leadership that creates an environment to encourage digital transformation will communicate a clear vision for digital government and actively champion its benefits. [Such] leaders will be engaged, visible and approachable, and empower their teams through decentralising decision making".[5] A study by the European Commission (EC) (2024[8]), based on a survey of 576 public managers in seven countries, found that leadership can especially influence AI adoption by offering robust incentives and/or financial resources to implement AI initiatives, with respondents generally finding the current state of these to be unsatisfactory.

Strong leadership for AI can foster a "mission-oriented" approach to its innovation. This approach emphasises a problem-solving focus, where policy interventions are designed to mobilise resources, coordinate stakeholders, and stimulate innovation and collaboration across government and sectors to tackle the identified challenge and meet set mission targets. Mission-oriented policies often involve a combination of regulatory measures, financial incentives, research funding and targeted investments to drive progress towards the mission. Leaders play a critical role by providing top-down direction and galvanising support in order to align all pieces of government to move in unison towards the same goal (OECD, 2021[9]).[6] Some governments and intergovernmental organisations have taken a mission-oriented approach to AI policy, although these are generally aimed at catalysing economic growth in the private sector (UCL IIPP, 2019[10]; Vinnova, 2022[11]).

Taking a strategic and directed approach

Governments can implement whole-of-government strategies and guidance on AI to identify and prioritise its coherent use and development in line with overarching government values and objectives. For example, the Dominican Republic's (2024[12]) 2023 national AI strategy focuses heavily on AI in government. Canada (2025[13]), Switzerland (2025[14]) and Uruguay (2021[15]) have each developed a dedicated strategy for AI in government, and another is under development in the United Kingdom (2024[16]). The April 2025 policy from the United States (US) on "Accelerating Federal Use of AI through Innovation, Governance, and Public Trust", although not called a "strategy", fits many of the hallmarks of a strategy that seeks to drive change throughout federal government and has the added benefit of being binding in nature (Box 4.1). Many others have substantive aims for AI in government to be embedded in broader national strategies. Overall, these high-level strategies tend to touch on many of the enablers, guardrails, engagement processes and types of use cases discussed in this report. Targeted strategies also exist to guide AI efforts in certain government functions. For instance, France has developed a strategy for using AI in HRM (see Chapter 5, Box 5.19), and the aforementioned US policy gives government agencies 180 days to develop and publish their own AI strategy.

Box 4.1. Accelerating federal government use of AI in the United States

In the United States, pursuant to requirements of the AI in Government Act of 2020, on 3 April 2025, the White House Office of Management and Budget (OMB) issued *M-25-21 Accelerating Federal Use of AI through Innovation, Governance, and Public Trust.* The policy promotes AI innovation, responsible adoption and use of AI, and safeguarding American protections on privacy, civil rights, and civil liberties. According to the policy, among other things, agencies must:

- Identify a Chief AI Officer (CAIO) to serve as a senior advisor, champion agency AI goals, coordinate AI efforts within in their agency, and represent the agency with coordination bodies and external fora; with OMB committing to convening an interagency CAIO Council to support coordination to maximise efficiencies.

- Remain accountable by meeting reporting requirements, including updating an AI use case inventory at least annually.

- Implement minimum risk management practices for AI that could have significant impacts when deployed ("high-impact AI") in a manner proportionate to the anticipated risk from its intended use, as described further (Box 1.3).

- Publish agency AI strategies for identifying and removing barriers to the responsible use of AI and achieving enterprise-wide improvements in the maturity of their applications (CFO Act agencies).

> o Strategies must include an assessment of the agency's current state of AI maturity and a plan to achieve the agency's AI maturity goals, by addressing plans and processes to, among other things, develop AI-enabling infrastructure (e.g. high-performance computing infrastructure) across the AI lifecycle; ensure access to quality data; develop enterprise capacity for AI innovation; recruit, hire, train, retain and empower an AI-ready workforce and achieve AI literacy for non-practitioners involved in AI; and develop the necessary operations, governance and infrastructure to manage risks from the use of AI.
>
> - Develop a GenAI policy that sets the terms for acceptable use of GenAI for their missions and establishes adequate safeguards and oversight mechanisms that enable GenAI to be used in the agency without posing undue risk.
>
> - Proactively share data and AI assets across the federal government, including custom developed code, including models, whether agency developed or procured, and to the extent practicable, release and maintain the code as open-source software in a public repository (some exceptions apply).
>
> - Develop and publish agency compliance plans to achieve consistency with M-25-21 (to be updated every two years).
>
> - Revisit and update where necessary internal policies on IT infrastructure, data, cybersecurity and privacy to align with M-25-21 and other relevant executive orders and laws.
>
> - Responsibly procure AI capabilities (Box 4.7).
>
> Note: The policy generally applies to all Executive Branch departments and agencies, including independent regulatory agencies. Some parts of the policy apply only to "CFO Act" (Chief Financial Officers Act of 1990) agencies. National intelligence agencies and the Department of Defense are also excluded from some requirements. The source document includes more specific details on applicability.
> Source: (US OMB, 2025[17]), https://www.congress.gov/bill/116th-congress/house-bill/2575/text, https://www.congress.gov/bill/116th-congress/house-bill/133.

In an attempt to move beyond strategy, some governments have developed comprehensive guidance. For instance, in addition to the AI Playbook for the UK Government (Box 4.2), New Zealand (2025[18]) has established a Public Service AI Framework, and Ireland (2025[19]) has published Guidelines for the Responsible Use of AI in the Public Service. Such guidance is helpful not only to implement strategies but can help overcome risk aversion in implementing more formal laws and regulations (see Guardrails discussion below) by removing the need for each organisation or team to make its own interpretations. As stated by the co-founder of Code for America, "well-meaning and well-written legislation originates at the top of a very tall hierarchy, and as it descends, the flexibility that its authors intended degrades. Laws often have an effect entirely different from what lawmakers intended because of this cascade of rigidity" (Pahlka, 2024[20]).

To further counteract risk aversion, guidance could promote experienced public servants' use of judgement and discretion. It could also acknowledge that, as with any human action, leveraging AI cannot be risk free. Techniques such as behavioural science can be used to craft guidance and communications in a way that helps ensure the desired behavioural effect on users, maximising the value of AI adoption while mitigating risks, increasing the likelihood of meaningful behavioural change and responsible AI adoption (OECD, 2021[21]).[7]

Box 4.2. Artificial Intelligence Playbook for the UK Government

The UK Government Digital Service (GDS), with the support of a variety of central government departments, private sector technology companies, academic institutions and users, published the Playbook in February 2025. It includes 10 principles alongside guidance on several issues.

10 principles to guide AI in government organisations

1. You know what AI is and what its limitations are.
2. You use AI lawfully, ethically and responsibly.
3. You know how to use AI securely.
4. You have meaningful human control at the right stage.
5. You understand how to manage the AI life cycle.
6. You use the right tool for the job.
7. You are open and collaborative.
8. You work with commercial colleagues from the start.
9. You have the skills and expertise needed to implement and use AI.
10. You use these principles alongside your organisation's policies and have the right assurance in place.

Explainers and guidance

The Playbook includes informational material on what AI is, fields of AI, applications of AI in government, limitations, building AI solutions, building a team, defining the goal, buying AI, using AI safely and responsibly, ethics, legal considerations, data protection and privacy, security and governance. It also includes a series of AI uses cases that illustrate real-world efforts in the UK government.

Source: (UK Government Digital Service, 2025[22]).

Determining whether AI is the best solution

Guidance should exist that includes a focus on determining whether AI is the best solution for a given problem, thus taking a step back from focusing on AI. While AI has tremendous capabilities, it is not always the best solution and in many cases is inviable. A key finding from a recent report from the Ada Lovelace Institute (2025[23]) is that "there is a surprising lack of evidence on the effectiveness and impact of AI tools, even from a purely technical standpoint. Evaluating AI interventions in context is crucial to determining their performance and value compared to existing manual or traditional methods". Work by RAND found that an inadequate of understanding of the problem to be solved and projects that use AI unnecessarily are two main drivers of AI project failure (2024[24]; 2025[25]).

A common issue with AI is that people start with solutions then look for problems for the technology to solve. In general, governments should seek to understand and focus on the outcomes both governments and citizens seek to achieve and the problems preventing that. Armed with this knowledge and priorities, they can then identify whether AI (or something else) is the best solution to help achieve these goals (Berryhill et al., 2019[4]; Mulgan, 2019[26]). Accordingly, governments need capacities for problem identification and understanding. They will also need to leverage other enablers presented in this chapter, including workforce skills for understanding AI's strengths and weaknesses relative to other technologies and processes to engage users to understand their needs.

Some governments have built in considerations for this in government-wide guidance. For instance, the AI Playbook for the UK Government indicates that public servants should "be open to the conclusion that, sometimes, AI is not the best solution for your problem: it may be more easily solved with more established technologies". The United Kingdom's (2020[27]) Guidelines for AI Procurement advises AI procurement officials to "start with [a] problem statement" and articulate "why you consider AI to be relevant to the problem, and be open to alternative solutions". In the United States, the AI Guide for Government ([28]) includes components to "focus on the root problem", and to consider "Is it the best option to solve this particular problem? Have you evaluated alternative solutions?"

A processes for making these determinations could be integrated as part of an ex ante impact evaluation (see Guardrails section below on "Impact assessments"), or could be established as an independent process before entering the pipeline of AI projects.

Defining clear roles and responsibilities

Governments should define clear roles and responsibilities to facilitate the coherent development, use and potential scaling of AI. These roles and responsibilities should be defined and agreed upon with relevant stakeholders and assigned to government institutions, incorporating them into the institutions' mandates. This enables a solid institutional structure to support the implementation of national strategies within and across individual institutions, and it facilitates accountability and oversight across the public administration. Several European countries are expanding the mandates of existing ministries or agencies to ensure coherent AI deployment. Examples include Norway's Ministry of Digitalisation and Public Governance, and Spain's Secretary of State for Digitalisation and AI within the Ministry of Digital Transformation and Public Management (OECD, 2024[3]). The United Kingdom has consolidated most AI roles and responsibilities across sectors into its Department for Science, Innovation and Technology (DSIT) (OECD/UNESCO, 2024[29]). At the individual level, the US policy in Box 4.1 requires federal government agencies to designate a CAIO to lead their AI efforts. Governments may also need to parse out various roles and responsibilities to achieve a variety of strategic objectives. For instance, in Chile and Colombia, the national AI strategies define responsible actors linked to each commitment as defined in the strategy. Colombia also details the time frames in which responsible actors need to achieve them, as well as budget and monitoring indicators (OECD/CAF, 2022[30]; CONPES, 2025[31]). In establishing roles and responsibilities, governments should make clear which entity or entities have the authority to establish policy over AI use in government. Indeed, OECD work with countries has uncovered confusion around who is responsible for rule-setting, potentially hindering digital transformation efforts.

Coordinating efforts within and across government

Governments can reinforce coordination and collaboration efforts to ensure a holistic approach to AI adoption and governance. Establishing inter-ministerial task forces or committees can facilitate decision-making, communication and collaboration across different institutions. These mechanisms enable all actors to take part in setting overarching objectives and work together towards achieving them. For instance, the US policy discussed in Box 4.1 requires the establishment of a cross-government CAIO Council. In another example, Australia, established a temporary AI in Government Taskforce (September 2023 to June 2024) to develop policy, standards and guidance to enable the safe, ethical and responsible use of AI in public service (OECD, 2024[3]).[8] At the sub-national level, in the government of Dubai in the United Arab Emirates, 22 CAIOs from across the government are charged with leading and coordinating AI efforts (WAM, 2024[32]).

Coordination is also important transversally, across levels of government. Many AI applications have significant local impact, particularly in public service delivery and social welfare, as local governments are closest to citizens and residents. However, without coordination, fragmentation in AI approaches can emerge. Where municipalities and regional governments develop AI solutions in isolation, inefficiencies,

duplication of efforts and inconsistencies in governance frameworks and user experience can result (Verhulst and Sloane, 2020[33]). Similarly, designing national strategies policies without consideration for local needs can result in approaches that are mismatched or unworkable for realities on the ground. Denmark has taken steps to address this challenge through its new Digital Taskforce for AI, which is working to scale AI adoption across all levels of government, ensuring alignment in priorities, standards and governance approaches.[9] Similarly, in Sweden, AI Sweden and Vinnova launched the Collaboration for AI in Municipalities and Civil Society initiative (Kraftsamlingen) to help municipalities and civil society organisations integrate AI into their operations. Since 2022, this initiative has provided tailored support, including guidance on AI adoption and funding opportunities for concrete projects, fostering a more coordinated and effective AI ecosystem across local governments.

Creating spaces to experiment

Governments need to allocate time and space to explore using AI, as both experimentation and iterative learning are crucial to developing AI capacity (OECD/CAF, 2022[30]). In addition to helping promote learning and identify new possibilities and approaches, controlled environments for AI experimentation and testing facilitate the timely identification of potential technical flaws, behavioural biases of both AI systems and people using them, and associated governance challenges. Furthermore, through experimentation, AI systems can be incubated until the solutions are technically robust enough to scale up. In doing so, they can also highlight public concerns especially through testing under quasi real-world conditions (OECD, 2019[34]). Such approaches entail engaging stakeholders throughout the development phase, evaluating user needs, assessing data availability and quality, and continuously monitoring progress from the prototyping and piloting phases (OECD/UNESCO, 2024[29]). Such environments include innovation centres, labs and sandboxes. Experiments can operate in "start-up mode" — whereby they are deployed, evaluated and modified — then be scaled up or down, or abandoned quickly (OECD, 2023[35]). Beyond in-house experimentation, governments can also work with non-governmental actors, such as GovTech startups, to design and execute AI experiments (see "Turning to GovTech startups" below).

Additionally, these environments foster collaboration between government, academia and industry, promoting the exchange of ideas and accelerating the development of AI technologies. By simulating real-world conditions, these facilities enable rigorous validation of AI systems, ensuring they are robust, reliable and safe prior to deployment. This approach to testing and experimentation not only enhances the effectiveness of AI solutions but also builds public trust by validating these technologies through early identification and addressing potential risks, biases or inefficiencies before wide deployment.

For example, the EU, in collaboration with its Member States, has launched a network of permanent Testing and Experimentation Facilities (TEFs), including CitCom.ai, which focuses on smart cities and communities. The initiative accelerates the development of trustworthy AI in the European Union by providing innovators — both companies and public agencies — access to test and experimentation of AI-based products in real-world conditions. Other examples include:

- In the United States, the Mitre Corporation, a government-funded research and development (R&D) centre, is developing an AI supercomputer to power a new AI sandbox, which will be capable of training new, government-specific advanced AI systems.[10]

- In the United Kingdom, the Incubator for AI (i.AI) promotes AI experimentation and eventual scale-up through four key approaches: 1) *prototyping* to quickly test and evaluate ideas for AI applications; 2) *delivery* to scale up successful prototypes to relevant government teams where they can have an impact; 3) *modularisation* to share technical work across government, including open sourcing code; and 4) *convening and advising* to identify areas for sharing learning and products.[11]

- In France, ALLiaNCE, an interministerial AI incubator launched by the Interministerial Directorate for Digital Affairs (DINUM) in July 2023, illustrates a government-led initiative to structure the

experimentation and scaling of AI across core government functions.[12] The ALLiaNCE incubator applies an agile, user-centric "product mode" methodology — originally developed by beta.gouv.fr — focusing on fast iteration, user feedback and measurable impact, thereby accelerating AI adoption in government. ALLiaNCE's structured selection criteria — based on impact, mutualisation potential, user engagement and ethical compliance — demonstrates a rigorous approach to testing and scaling responsible AI projects.

- In Australia (2024[36]), over 60 Australian Public Service (APS) agencies conducted a six-month trial of Microsoft Copilot. Over 7 700 public servants participated in the trial, the results of which were varied; however, aggregately, users experienced perceived improvements in the efficiency and quality of AI for summarisation and preparing a first draft of documents. Despite this, AI's adoption requires concerted efforts to address technical, cultural and capability barriers.

- Portugal and Spain have each engaged with GovTech organisations to promote experimentation with digital technologies, including AI, in justice administration.[13] Also on GovTech but with a broader scope, Spain's GovTechLab is an AI use case incubator that identifies scenarios where generative AI can have an impact on public administrations – whether by achieving greater efficiency in the provision of public services, reducing workloads or improving citizen service.[14] Twenty out of the 300 identified use cases will be piloted in areas such as document classification, AI assistants and the preparation of tenders and grants. Those that are successful will be scaled and offered as a service to the entire administration.

In establishing small AI experiments and pilots, governments should consider their definition of success and establish measurements and evaluation framework to determine whether a project was successful, determine what worked and what did not and help capture and disseminate lessons learned. The United Kingdom's Guidance on the Impact Evaluation of AI Interventions serves as a good example, providing considerations on AI projects from small-scale testing through full implementation.[15]

The OECD is currently developing a dedicated report on AI experimentation in government to review current practices and derive key lessons learned to help inform and guide policymakers in establishing their own experimentation guidance for their organisations (OECD, forthcoming[37]).

Creating a strong data foundation

Data serves as the foundational asset driving the capacity of AI to function, evolve and create public value. Drawing on the concept of "garbage in, garbage out," AI performance directly correlates to the quality and representativeness of inputs it is trained with; AI systems often require vast amounts of data across the AI system cycle to deliver valuable outputs.

Relevant data can be derived from government, the private sector or other sources. This section mainly focuses on government data, though the OECD (forthcoming[38]) is conducting work that systematises the sources from which AI developers obtain data for AI training and highlights their main attributes.

Access to and sharing government data for AI brings complex data governance challenges. Governments face regulatory and operational hurdles, from safeguarding privacy, non-biased results and data security to navigating policy and legal frameworks for governing data sharing and use and intellectual property rights. Public sector organisations also have to tackle technical issues, from ensuring interoperability across data systems to building the technical capacity to manage data effectively.

Ensuring privacy, security and intellectual property rights

Governments are building frameworks, guidelines and mechanisms to promote strong data governance that safeguard privacy, intellectual property and security. These often result from collaborations among regulatory bodies, industry stakeholders and civil society. For example:

- Korea's Personal Information Protection Commission (2023[39]) has released a guide for personal information processing and AI development. This guide outlines the legal basis for processing personal data, establishes safety standards, and suggests measures to protect individuals' rights within AI systems.

- New Zealand's (2025[18]) Public Service AI Framework and the United Kingdom's (2025[22]) AI Playbook for the UK Government each cover principles for safe and privacy-preserving use of AI in government.

- In bridging the concepts of experimentation (discussed above) with personal data protection, France has established a data sandbox to provide an enabling environment for safe experimentation, coupled with training and hands-on support in managing personal data and ensuring regulatory compliance (Box 4.3).

- Some governments are exploring privacy-enhancing technologies (PETs), such as data anonymisation, on sensitive data used for AI training. These technologies can, in turn, be enhanced through AI (OECD, 2024[40]).[16]

Box 4.3. France's personal data sandbox for AI in public services

In 2023, France's data protection authority (CNIL) launched a "sandbox" initiative to support innovation in AI for public services. This sandbox offers selected organisations expert guidance to help them navigate personal data regulations early in their project development. While it does not remove any legal requirements, it aids in identifying solutions to compliance challenges. Four projects were selected to receive guidance from CNIL's new AI department:

- Albert (DINUM): assists civil service agents with a language model to improve responses to user inquiries, with pilots in "France Services" centres.
- Job Intelligence (Pôle Emploi): provides personalised job search guidance using professional data to match job seekers with tailored services.
- Ekonom AI (Nantes Métropole): offers water consumption insights and recommendations for residents, supporting ecological goals and potentially adaptable to other public policies.
- RATP Video Project: develops AI to detect events through matrix data capture, ensuring privacy by design with no personal data collection.

Source: (CNIL, 2023[41]; [42]).

Ensuring representativeness in data

Ensuring AI systems are trained on representative data is crucial for delivering accurate and relevant outcomes. In some countries, different populations have unique languages and traditions. In others, different demographic or other contextual factors shape the data needed for AI to be effective. As described in the OECD Good Practice Principles for Data Ethics in the Public Sector (2021[43]), using data that is not representative to train AI can lead to significant issues, particularly for government applications that require fair and accurate policies and decisions that can tangibly impact the target population. These issues include biased algorithms and decisions, and an inability to develop tailored services and policies for groups underrepresented in data, as discussed in Chapter 1.

Governments are taking actions to address this issue. For instance, a number of countries have taken action to invest in efforts to promote representativeness in languages (OECD, 2023[44]; Peixoto, Canuto and Jordan, 2024[45])[17]. Examples include:

- To promote language representation, the Common Danish Language Resource initiative and the Danish Platform for Danish Language Resources, led by the Danish Agency for Digital Government (2024[46]), aim to collect, develop and display language data and other tools that can support the development of Danish AI solutions.

- Greece (2024[47]) is pursuing the development of a Greek Language and Culture Data Space, which focuses on integrating Greece's linguistic and cultural heritage into AI applications.[18] Also related to the Greek language, the development of the "Meltemi" and "Llama-Krikri" LLMs represent promising steps in this direction, highlighting the importance of open access and collaborative efforts in expanding the available linguistic resources.[19]

- India's Bhashini platform, launched under the National Language Translation Mission (NLTM), is India's flagship AI-led language infrastructure project. It supports real-time translation across 22 official Indian languages and dozens of dialects and provides support for multilingual voice assistants and AI service delivery interfaces.[20] It was made possible by through a massive citizen engagement exercise to build multilingual datasets that boost data representativeness, with models and application programming interfaces (APIs) made available as open source.

- In Saudi Arabia, the Saudi Data and Artificial Intelligence Authority (SDAIA) has launched ALLaM, an LLM developed with 500 billion tokens and over 300 000 Arabic texts, including encyclopaedias, scientific research and historical works (M Saiful Bari, 2024[48]). ALLaM aims to reflect the linguistic and cultural richness of the Arabic language.

- Spain is working on a family of AI models, called ALIA, that are heavily trained on native Spanish and other official language data and will be available as open source.[21] The first use cases include an assistant for diagnosing heart failure in the public health sector and an assistant to facilitate tax officials' replies to citizens.

- Language models for native American languages have also been developed. Although not developed by government, researchers have developed LakotaBERT to support language revitalisation efforts for "Lakota, a critically endangered language of the Sioux people in North America" (Parankusham, Rizk and Santosh, 2025[49]).

Enabling effective and trusted data access and sharing

As discussed in Chapter 3, governments often face a significant shortage of easily available, relevant, high-quality data necessary for training AI systems effectively (OECD, 2025[50]). Addressing this gap requires a focused effort on enhancing data access, including through collaborative data collection and open arrangements (Box 4.4).

Box 4.4. Sweden's collaborative data gathering for Svea

Svea is a Swedish initiative coordinated by AI Sweden that unites government agencies, municipalities, regions and industry to address the challenges of creating AI solutions for public services. The primary focus is pooling resources to gather Swedish-language data that reflects the unique needs of government — a task too large for any single organisation.

By collaborating, government organisations can share the workload of data collection essential for developing a useful AI assistant. In the first phase, participants identified specific needs and began generating data from within their organisations to train the system. In the upcoming phase, they will gain access to shared databases of relevant national information to further inform the AI assistant.

Source: (AI Sweden, 2024[51]).

Another key initiative is open government data. On average, of only 46% of high-value government datasets are available as open data across the OECD (2023[52]), compared to more than 80% in France and Korea. Challenges remain in terms of fostering open data re-use by actors and the integration of these datasets into AI systems. Results from the 2023 edition of the Open, Useful and Re-usable data (OURdata) Index show that countries perform better in data availability and data accessibility compared to government support for data re-use (Figure 4.1).

Figure 4.1. OECD Open, Useful and Re-usable data (OURdata) Index, 2023
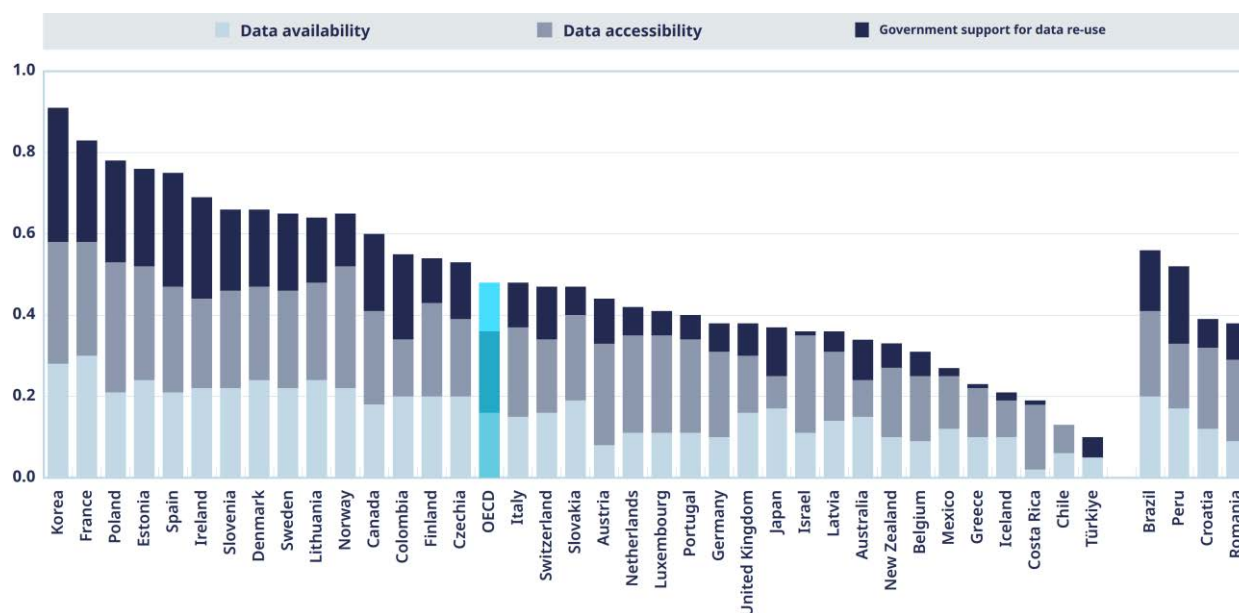
Source: (OECD, 2023[52]).

A key question is how to increase the value of open government data for AI systems by design, and thus its accessibility and AI-readiness. On the one hand, the standardisation (e.g. in terms of structure and formats) of open government data can reduce the time AI-developers need to invest in preparing data to train AI-systems. On the other hand, the increased use of tools, such as APIs, can also support data integration with AI-systems by providing a standardised method for sharing and accessing data automatically, directly from its source. Today only 47% of high-value datasets are released with APIs (OECD, 2023[52]).

Beyond open data, other relevant initiatives include increasing access to large government-held or publicly funded datasets, such as Korea's AI Hub Data Finder (2024[53]), which provides access to text, imagery, video, audio and sensor datasets relevant for AI-training in areas such as healthcare, public transportation and disaster and safety.

Using private sector data

Both public and private sector data can play a vital role in developing AI applications for government. While government data provides essential insights into demographics and public services, private sector data — such as on mobility patterns, consumer behaviour and financial trends — can enhance these insights. For example, AI systems for urban planning can benefit from telecommunications data to analyse traffic flow, while healthcare AI can leverage anonymised patient data from private clinics to improve disease prediction. By combining both sources responsibly and ethically, AI applications can become more accurate, efficient and responsive to public needs. One example of pooling and combining data from the public and private sectors is the Common European Data Spaces. The purpose of the data spaces is to

make more data available for access and re-use across the European Union in a trustworthy and secure environment for the benefit of European businesses and citizens (OECD, 2024[54]; EC, 2025[55]).

Creating a conducive environment with data governance

From enabling data access and sharing to building the foundations necessary to make AI in government a possibility, governments need to develop robust "data governance" arrangements, which can be integrated into broader AI strategies and policies (OECD, 2024[3]).

> **Data governance** *refers to "diverse arrangements, including technical, policy, regulatory and institutional provisions, that affect data and their creation, collection, storage, use, protection, access, sharing and deletion, including across policy domains and organisational and national borders" (OECD, 2022[56]).*

Originally developed to explore the specific arrangements that should be in place to enable data access and sharing, the OECD framework for data governance in the public sector can be applied to the context of data for AI-systems (Figure 4.2). Governments can develop data governance capabilities in the public sector by prioritising the development of comprehensive data strategies, defining leadership roles and establishing a vision for managing and governing data at a more technical level to realise AI's intended benefits and outcomes.

Figure 4.2. Data governance in the public sector
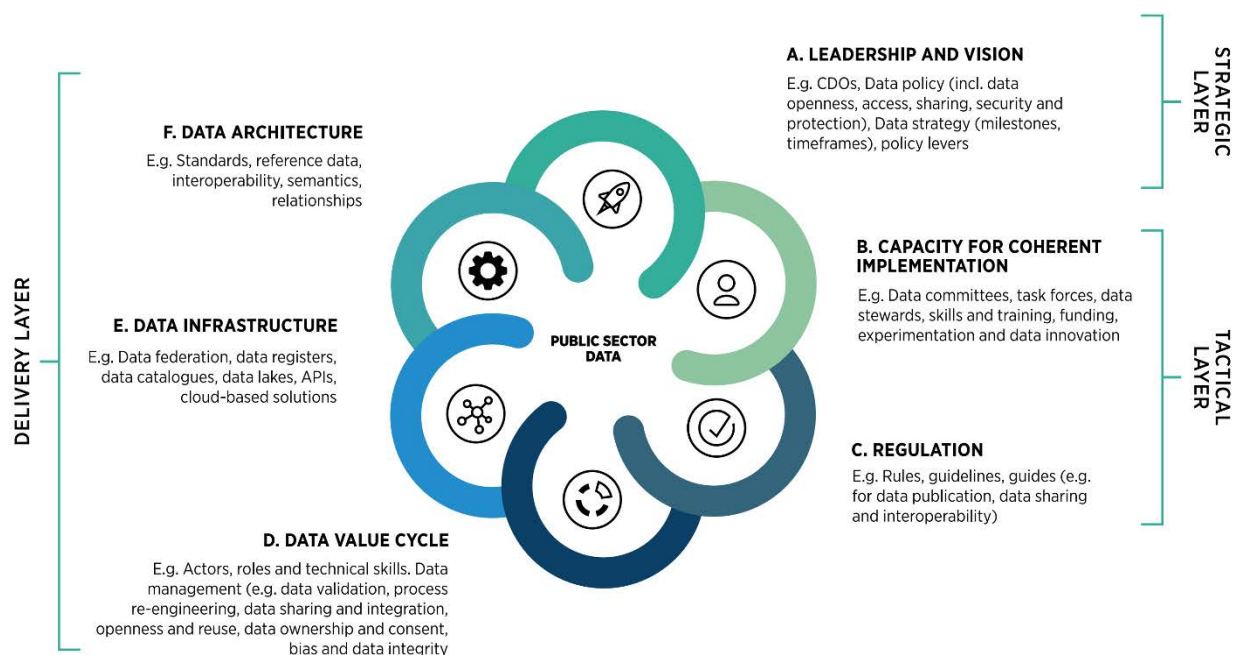
Source: (OECD, 2019[57]).

For example, Canada's 2023-2026 Data Strategy for the Federal Public Service outlines the desired outcomes and guiding principles to advance sound data governance across the federal government as a whole, along with expectations for roles and responsibilities.[22] Other examples include the United Kingdom's National Data Strategy and Australia's Data and Digital Government Strategy.[23] Greece (2024[47]) is pursuing a flagship programme on Data Governance and AI Strategy Coordination to establishing elements necessary to support an AI-ready public and private sector. The US policy discussed in Box 4.1 puts in place several requirements to improve data governance. It is also important to embed stakeholder engagement with data rights holders, such as with citizens, businesses and civil society representatives, who could be impacted by the use of data for AI — either in the context of intellectual

property rights or personal data rights (OECD, 2022[58]), or by the potential use of inadequate or skewed data.

Governments should support their vision and strategy on data for AI with adequate capacity for consistent implementation across the public administration, along with guidelines and legal frameworks to ensure effectiveness (OECD, 2019[59]). This can include numerous areas, such as improving data literacy and skills for AI, with examples such as Brazil's CAPACITA GOV.BR initiative[24] and Argentina's National Programme for Enhancing the Protection of Personal Data. This can also include efforts to boost coordination and institutional collaboration, with examples such as the Norwegian Resource Centre for Sharing and Use of Data (Digdir, 2024[60]). Finally, it is essential to recognise the enabling role of legal frameworks that orchestrate and accelerate the integration and exchange of data between public institutions, while safeguarding individual rights and privacy. Clear, modern and applicable legislation on data governance and personal data protection is important for deploying trustworthy AI systems at scale. In Chile, a draft Data Governance Bill is currently under discussion, which defines principles, roles and interinstitutional coordination mechanisms for data governance. In addition, a new Personal Data Protection Law, aligned with international standards, is entering into force. Other examples include Ireland's Data Sharing and Governance Act 2019 and the EU General Data Protection Regulation (GDPR), Data Act, Data Governance Act, Open Data Directive and data interoperability frameworks.[25]

Finally, the delivery portion of data governance refers to the processes, mechanisms and tools that enable the operational implementation of data governance at the organisational and team level, ensuring that sound data governance and management practices are implemented and integrated across the AI data value cycle (OECD, 2019[59]). One example is the United States (2024[61]) assessment for Data Operations (DataOps) maturity across federal agencies as part of its AI Guide for Government. This framework evaluates how well organisations can discover, access and utilise data to support AI development throughout the data value lifecycle. Key components include securing a comprehensive data asset catalogue, flexible data access methods and tools that facilitate documented AI experiments. Other key components of delivering data governance are the technical skills and job profiles needed, including data scientists, domain experts, data engineers and data providers, who are involved in data collection and processing for AI (OECD, 2022[58]). Issues related to data infrastructure are covered in the next section.

Building out digital infrastructure

Ensuring the availability of reliable and scalable digital infrastructure can assist in supporting and scaling AI in government. In addition to data itself, as discussed in the previous section, data infrastructure, scalable computing platforms, AI foundation models and common AI tools are important building blocks for AI in government. Other forms of digital infrastructure also exist and are discussed in OECD work (2024[62]). This section focuses on those most relevant to AI.

In their attempts to purse trustworthy and scalable AI systems, governments are confronted with strategic decisions with regards to AI systems. On one hand, building national capacity through the development of national digital infrastructure can help a country in implementing its own data protection and privacy rules. A number of countries are seeking to develop such capacity (Letzing, 2024[63]; France Élysée, 2025[64]; African Union, 2024[65]; Ray, 2025[66]; EC, 2025[67]; Brizuela et al., 2025[68]). On the other hand, this could also contribute to technological fragmentation and closed ecosystem that limit international collaboration (Komaitis, Ponce de León and Thibaut, 2024[69]; Frazier, 2025[70]). Governments need to consider various options in determining a balance that is appropriate to them for developing solutions in-house versus in collaboration with the private sector and with other countries.

Digital infrastructure is not only relevant for national governments; it can also be formative in AI adoption in subnational governments, such as cities. The development of shared and reusable digital tools can help local governments overcome the entry costs due to the underlying economies of scale and allow AI solutions to be tailored to local needs and contexts.

Computing power and data infrastructure

Access to computing infrastructure resources can be key to the effective development and use of AI in government (OECD, 2022[71]). Choosing between on-premises and cloud solutions for AI deployment depends on specific needs, political choices, regulatory requirements, budget constraints and long-term goals.

On-premises solutions offer greater control, customisation and security, making them suitable for highly sensitive applications and to conform with data localisation laws (Redapt, 2023[72]). Cloud solutions, on the other hand, provide unparalleled scalability, cost efficiency and access to cutting-edge AI technologies, making them ideal for dynamic and rapidly evolving projects and more practical than on-premises solutions for small projects or newer or smaller entrants to AI development (Dombo, 2023[73]). More than half of OECD countries have cloud technology initiatives in place, including storage and computing capabilities (Infrastructure as a Service, or IaaS). Notably, access to cloud technologies relies on both public and private solutions (48% vs. 52% respectively), with several countries pursuing the development of public-sector-led cloud technologies (OECD, 2024[74]). In many cases, a hybrid approach combining both on-premises or otherwise dedicated infrastructure and public cloud (shared, third-party) resources ("hybrid cloud") can offer a balanced solution, leveraging the strengths of each. Notably, work by RAND (2024[24]; 2025[25]) found that among companies, those who are able to use cloud solutions generally did not face challenges in securing adequate compute, but those who could not transfer data to the cloud faced significant challenges that contribute to AI project failure.

The global demand for AI-ready data centre capacity could triple by 2030, demonstrating the growing use of AI (McKinsey, 2024[75]). Carbon emissions produced by data centres and data transmission networks are already estimated to be 1% of all energy-related emissions, but have generally grown only modestly despite rapidly growing demand for digital services due in part to increased hardware and model efficiencies over time (OECD, 2022[71]; IEA, 2023[76]). Recent analysis from the International Energy Agency (IEA) (2025[77]) found that data centres are among the fastest growing sources of emissions and that such emissions could increase significantly in the next ten years, but also that "widespread adoption of existing AI applications could lead to emissions reductions that are far larger than emissions from data centres". Nevertheless, the sizeable carbon emissions linked to data highlights the need for practices to manage AI's energy requirements. Increasing amounts of water needed to cool data centres is also important to recognise (Metz et al., 2025[78]). Recent research and industry developments indicate a growing trend in AI towards the adoption of smaller and/or more specialised models, such as small language models (SLMs) that consume fewer resources, require less data and are less expensive (Hassani et al., 2022[79]; Jones, 2025[80]).

> **Box 4.5. Korea's shared data centres and government cloud**
>
> Korea's National Information Resources Service (NIRS) has been working with the Ministry of the Interior and Safety to upgrade key hardware, networks and management tools to help modernise Korea's technology and enable migration to the cloud. A critical part of this has been the construction of new government data centres, which can help ensure compliance with government requirements, cost-efficiencies with a reduced technology footprint, and job creation and local investment in target areas. These data centres have also been made available to the government's main partners in the private sector, which helps ensure that companies holding or handling sensitive data are doing so in an environment that meets the government's requirements for security, back-up and redundancy, among others. With measures around sustainability and renewable energy, the data centres help reduce the environmental impact of Korea's digital government, particularly as it prepares to make greater use of AI solutions.
>
> Source: OECD Digital Government Review of Korea (forthcoming).

In seeking a holistic approach, some countries are establishing compute and data infrastructure as part of a package of efforts, including approaches to digital public infrastructure (DPI),[26] to support the secure and meaningful exchange of data across government, underpinned by strong data integration and analytics capabilities. Such infrastructures not only help scale AI use but can also foster interinstitutional collaboration and the generation of public value from data. Take, for example, Brazil's National Data Infrastructure (IND).[27] This strategic initiative establishes a set of policies, standards, technologies and governance mechanisms to organise, share and manage public sector data securely and efficiently. Its main objective is to make government data findable, accessible, interoperable and reusable (FAIR principles), promoting transparency, the improvement of public services, administrative efficiency and evidence-based decision-making throughout government, serving as a foundation for digital transformation and innovation. The country's gov.br platform serves as a central hub for integrating access to nearly 5 000 digital public services and includes a Conecta gov.br platform as a data interoperability layer across government. These platforms and other digital infrastructure are a foundation for AI in government. In another example, Saudi Arabia's national cloud platform, Deem Cloud, developed by the Saudi Data and Artificial Intelligence Authority (SDAIA), consolidates digital infrastructure across more than 190 public entities and over 260 data centres (SDAIA, 2025[81]). It provides a suite of cloud services to support secure and efficient digital operations, and has contributed to energy and cost savings, as part of broader efforts to modernise public sector infrastructure and support national digital strategies. Focused more on compute, Greece (2025[82]) is building DAEDALUS, which is set to be one of the most powerful supercomputers in Europe and will be accessible to public institutions.

Developing AI foundation models

Foundation models are a form of AI models trained on large amounts of data — generally using self-supervision at scale — that can be adapted to a wide range of downstream tasks (OECD, 2024[83]). Governments can develop their own foundation models or build upon existing ones to create approaches tailored to the specific context of a country and/or its public. Foundation models can be "fine-tuned" through further training on narrower datasets related to a particular task or domain, enhancing its performance for that specific context (Montgomery, Rossi and New, 2023[84]).

Building a clean sheet foundation model is generally considered expensive and typically requires significant data and power resources. Examples of privately developed, proprietary foundation models include Mistral Large and those that power Anthropic's Claude (e.g. Claude 3.7 Sonnet), Google Gemini (e.g. Gemini Ultra) and OpenAI's ChatGPT (e.g. GPT-5).

Although they can be costly and require vast data assets and energy use, governments can indeed train and build their own foundation models, as discussed in Chapter 3. Governments can also fine-tune and tailor a proprietary foundation model to better meet its own context, which can significantly reduce the financial and time costs associated with deploying AI for specific tasks. An example of this approach is Portugal's ChatGPT-enabled public virtual assistant for public services (Box 5.46).

Governments can also use "pre-trained" open-source models. These are foundation models that have been trained by a company or other organisation that made its "model architecture and weights freely and publicly accessible for anyone to modify, study, build on and use" (Seger et al., 2024[85]).[28] Most open-source models are created by large technology companies, such as Meta's Llama series, although more organic, open-source community-driven models have been developed, such as the BigScience Large Open-science Open-access Multilingual Language Model (BLOOM).[29] An example of a government leveraging open-source AI models is France's Albert virtual assistant for public servants (Box 5.46).

A foundation model that is tailored to a national and governmental context — such as through fine-tuning a proprietary model or customising an open-source model — can significantly reduce the cost of adoption for teams wanting to deploy AI. Yet development and use of foundation models comes with risks that governments should keep in mind, as discussed in Chapter 1.

Governments are increasingly demonstrating interest in investing in national or regional foundation models to enhance technological sovereignty and better reflect different languages and cultures. For instance, Latam-GPT, led by Chile, is being developed by over 30 Latin American institutions to create a model trained on regional data (Gob.cl, 2025[86]), OpenEuroLLM is an EU-funded initiative to build open-source models covering all official European languages (EC, 2025[87]), and context and language-tailored models have emerged in Southeast Asia (Noor and Kanitroj, 2025[88]). Similarly, Italy's Minerva project has developed the first LLM trained from scratch for the Italian language and is one of the few examples in this context of tailored foundation models being developed from the ground-up (University of Rome Sapienza, 2024[89]).

Common AI tools

Common AI tools that can be used across an entire government, and tailored to specific governmental needs, can serve as a form of DPI that enables and enhances other services. Sometimes built with foundation models — sometimes provided as another type of DPI — these tools provide a shared service layer and can support the automation of routine tasks, improve user interaction and enhance service delivery.

For instance, chatbots can handle a large volume of citizen inquiries, providing instant responses to common questions and relinquish human resources for more complex tasks. This not only improves efficiency but helps ensure that public services are more accessible and responsive to the needs of the community. To be considered DPI, these AI tools should solve a common, basic need and thus be usable across a wide range of public sector organisations. An example of this approach is Singapore's Virtual Intelligent Chat Assistant (VICA), provided as a shared service and used by more than 60 government agencies to create over 100 chatbots.[30]

Common tools for supporting AI do not always use AI. In other examples:

- Singapore's Whole of Government Application Analytics (WOGAA) has been developed as a government tool for monitoring the performance of government websites and digital services, including those enabled by AI, providing a central dashboard to track website traffic, automated reports with key metrics, benchmark performance compared to other government websites and more.

- In Estonia, to implement the once-only principle, all public databases are mandatorily described in the catalogue of interoperability resources (RIHA), which serves as the national registry of systems,

components, services, data models, semantic assets and more, guaranteeing the transparent, balanced and efficient management of public information systems.

- France's aforementioned ALLiaNCE AI incubator fosters the development of reusable AI products across administrations, aiming to mutualise efforts and reduce duplication. ALLiaNCE provides a multi-layered service offering, including AI-enabled tools embedded in France's digital suite (La Suite Numérique), as well as a foundational layer — Albert API — to support cross-sectoral re-use of GenAI systems. As a digital common, Albert API provides open, reusable GenAI systems, lowering the adoption threshold for public agencies and contributing to a national DPI for AI. The ALLiaNCE incubator also prioritises data sovereignty and open, sovereign digital solutions, reflecting government efforts to mitigate dependency risks and ensure trust in AI systems.

The OECD.AI Catalogue of Tools & Metrics for Trustworthy AI includes a variety of tools from inside and outside governments that may help inform them in determining their own needs.[31]

Fostering skills and talent

As discussed in Chapter 3, skills gaps are one of the most significant challenges to the adoption of trustworthy AI in government. Governments should therefore take commensurate actions to build internal competency and capacity.

Governments should equip civil servants with the right skills to maximise AI's effectiveness, while ensuring safe, secure and trustworthy use. An AI-ready public service is instrumental for the development and deployment of AI solutions, as well as for the effective use of AI-powered tools to enhance daily tasks and policymaking. A strategic and coordinated approach to AI skills and talent can help target different groups within the workforce, identify skills gaps, develop the right skills and attract and retain more specialised AI talent. Strong in-house skills can also contribute to building national capacity, a topic discussed in the previous section.

This may include recruiting individuals with the skills needed to work with AI, as well as the upskilling of existing roles. Governments will need to anticipate that as AI evolves, the necessary skills will also, calling for continuous learning. A solid approach will require a needs assessment to map the current level of capability for data and AI in the existing workforce, to identify the key gaps and inform the strategy to addresses these needs. This would mean being able to take informed decisions among recruitment, retention and development of digital talent. This could also inform tailored training programmes to address skills gaps, and an approach to manage and train (i.e. through skilling, upskilling or reskilling) the roles most affected by the integration of AI.

Assessing the needs of different user groups

Government institutions should assess the needs of different user groups to take up AI and use it effectively. An AI-ready workforce ranges from general users of AI systems to institutional leaders, data and digital professionals, and the more specialist roles. As illustrated in Figure 4.3, user groups become narrower and more specialised further down the pyramid.

Figure 4.3. Considering the level of AI literacy needed for different user groups in the workforce
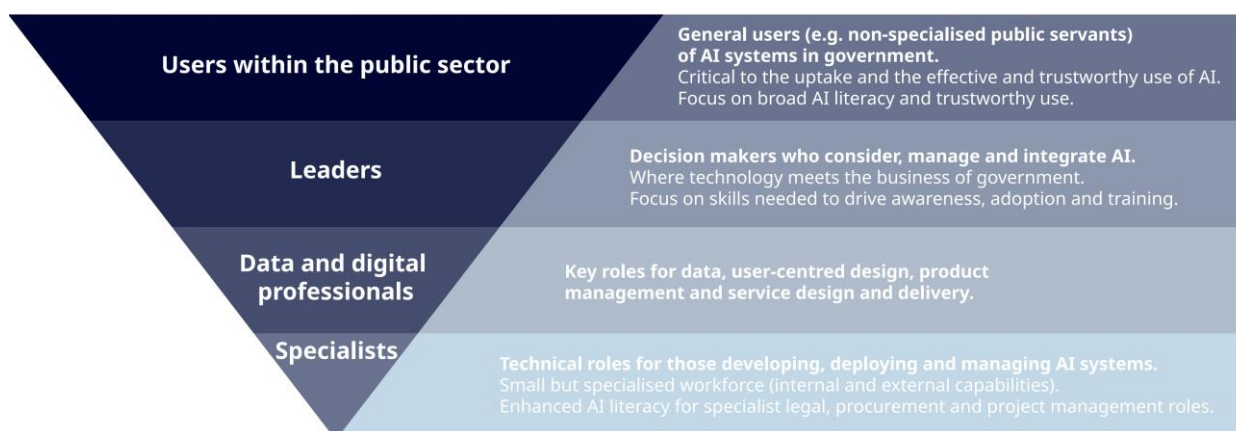
General, non-specialised public servants are critical to the adoption and effective use of AI. Their training should focus on general literacy around data and AI technologies, their effective use with respect to given tasks and the ethical and legal consideration for their use.

Leaders will be integral to the adoption of AI, being the layer where the technology meets the business of government. They help raise awareness among users, drive adoption of AI and promote learning and development opportunities in public administration. This user group needs a strategic vision and executive-level understanding of what AI technologies can do, their impact and how to address risks, compliance, funding and workforce management.

Data and digital professionals spearhead and facilitate the design, development and implementation of specific services. This user group needs a higher degree of AI literacy to comprehend how AI should be deployed to deliver the intended objectives for the services they are responsible for. Along with specialists, this group may be responsible for procuring AI through public procurement processes. Thus, the right set of skills can empower them in negotiations with vendors seeking to sell AI products and services.

While a smaller part of the workforce, **AI specialists** are critical for the development, deployment, management and use of AI systems. They extend beyond direct AI development, including roles in procurement, legal and project management. To develop this user group, governments need to target attraction, retention and learning and development. Additionally, to overcome potential skill shortages in the market, government organisations can also consider leveraging external capabilities through public procurement and partnerships, as discussed below.

Preparing civil service users

A civil service ready for AI will require a combination of foundational digital skills and more specific literacy in data and AI. The *OECD Policy Framework for Digital Talent and Skills in the Public Sector* (OECD, 2021[7]) outlines the various foundational digital skills applicable to every public servant, essential for supporting digital transformation:

- understanding potential of digital transformation
- understanding users and their needs
- collaborating openly for iterative delivery
- trustworthy use of data and technology
- skills to enable a data-driven public sector
- digital government socio-emotional skills
- digital government leadership skills.

For an AI-ready workforce, it is necessary to build on these foundations with the literacy needed for "individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool online, at home, and in the workplace" (Long and Magerko, 2020[90]). This includes understanding AI systems, data handling and management, and ethics. One approach, for example, is the AI Skills for Business Competency Policy Framework developed by The Alan Turing Institute in the United Kingdom (2023[91]), which has five dimensions:

- **Privacy and stewardship:** to mitigate risks around data security and protection, especially with regards to legal, regulatory and ethical considerations.
- **Specification, acquisition, engineering, architecture, storage and curation:** for the handling and management of data to enable more effective and ethical use of AI systems.
- **Problem definition and communication:** to identify, define and communicate those 'problems' that benefit most from the application of AI solutions.
- **Problem solving, analysis, modelling, visualisation:** including range of tools and methods that can be used for analysis, AI application and communication.
- **Evaluation and reflection:** to understand the impact of work on AI, assess efficiency and effectiveness of AI projects and identify opportunities to improve.

The level of competency and the extent to which each is required varies depending on the user group, existing capability or level of capability required for their roles. However, this combination of digital foundations and AI skills should lead to a more effective use of AI in government.

Developing skills and talent

To develop AI skills and talent in the public service, governments should consider both internal development practices, as well as the broader recruitment of key talent. Key mechanisms for developing existing talent include the items below, with additional examples provided in the OECD/UNESCO (2024[29]) G7 Toolkit for AI in the Public Sector.

- **Competency frameworks** outline the key skills and learning pathways for the workforce, which should be tailored according to the needs assessments and user groups outlined above. This could also contribute to the professionalisation of key AI roles for government. For example, the EC's Joint Research Centre (JRC) has developed a comprehensive competency framework (Box 4.6). India has also developed a dedicated competency framework to prepare public officials to lead AI transformation responsibly, tailoring content for different types of officials.[32]
- **Formal learning** can include courses, workshops and online modules. For example, Ireland provides a number of relevant courses, including training on AI in the Public Service and on the country's Guidelines for the Responsible Use of AI in the Public Service; and Greece's Hellenic Ministry of Internal Affairs in collaboration with Google has created AI training courses for public servants.[33] The globally available free and open course Elements of AI can help improve foundational AI literacy for public servants and citizens alike.[34]
- **Informal learning,** including through communities of practice, mentoring or job rotations, among others, can help raise awareness and drive adoption. Such approaches are discussed further below.

Box 4.6. European Union competency framework for AI in government

The EU has developed a comprehensive competency framework to guide public servants in effectively adopting and managing AI. This framework, elaborated by the Joint Research Centre (JRC), is based

---

on empirical research, including literature reviews, expert workshops and case studies, and identifies key competencies necessary for AI integration in public administration.

The framework classifies competencies into three main dimensions:

- technical competencies: encompassing knowledge and skills related to data management, machine learning and AI system implementation
- managerial competencies: addressing project ownership, knowledge brokering and decision-making in AI-related initiatives
- policy, legal and ethical competencies: covering AI procurement literacy, auditing and collaboration with domain experts to ensure compliance and ethical considerations.

Additionally, competencies are grouped into three cross-cutting clusters related to (i) attitudinal competencies (know-why), referring to mindsets and dispositions that support AI adoption, such as technology inquisitiveness and a data-oriented culture; (ii) operational competencies (know-how), covering practical skills for AI implementation, including database management, algorithm training and decision-making processes; and (iii) literacy competencies (know-what), related to fact-based knowledge of AI concepts, regulatory frameworks and fundamental machine learning principles.

Beyond developing the framework, the JRC's work gives three key recommendations on competencies:

1. develop focused, interdisciplinary AI competence training programmes
2. promote applied interdisciplinary research on AI competences
3. establish dedicated hiring processes and devote additional resources to attracting specialists with AI competences.

Source: https://publications.jrc.ec.europa.eu/repository/handle/JRC138702.

Developing AI-related skills internally should be complemented by external strategies to attract top talent and retain skilled public servants. The European Union's competency framework for AI in government highlights the need for a structured approach to AI skills development, emphasising technical, managerial and policy-related competencies. Canada's *Digital Talent Strategy* aligns with these principles, recognising that AI adoption in the public sector requires a balance of attitudinal (know-why), operational (know-how) and literacy (know-what) competencies to build an AI-ready workforce.[35]

To address critical skill gaps, targeted recruitment efforts and dedicated hiring processes are essential. The US Office of Personnel Management's (OPM) Skills-Based Hiring Guidance and Competency Model for AI offers a structured approach to defining and assessing AI job classifications,[36] while the EU framework underscores the importance of interdisciplinary training programmes and applied research. Governments should also focus on competitive compensation, clear career pathways and workplace flexibility to attract and retain AI talent.

Where internal capacity remains limited, partnerships with industry and academia, as well as strategic procurement of external expertise, can provide necessary support. The sections below on public procurement and AI partnerships explore these strategies further, offering policy examples. Additionally, continuous assessment of workforce alignment, progress in closing AI skill gaps and the effectiveness of learning initiatives is critical. By integrating competency-based insights from global best practices, governments can ensure their digital workforce evolves alongside technological advancements, ethical considerations and shifting labour market demands.

Facilitating connections and knowledge exchange

Communities of practice and networks allow for collaboration, learning, the sharing of expertise across organisational boundaries and the identification of collective or common problems. They can also serve as a useful conduit for soliciting user feedback on internal AI systems and services. The EC JRC competency framework (Box 4.6) notes that such activities can be critical for helping public servants gain know-how on AI and to help overcome challenges with early AI adoption. This framework outlines three action points for developing communities of practice: 1) create associations with relevant stakeholders, 2) deploy digital platforms for the communication and collaboration of involved entities, and 3) finance synergy grants for public-private collaboration and knowledge exchange, which can help reduce knowledge asymmetries and spark joint ventures.

Disseminating successful methods, strategies and use cases through such methods can help government organisations replicate and scale AI projects more effectively (OECD/UNESCO, 2024[29]). This approach helps avoid common mistakes, helps ensure consistency and accelerates the adoption of AI technologies across various government entities. For example:

- Estonia's skills development programme includes a data expert network with 500+ participants, AI meetups and experimentation events (e.g. hackathons, competitions) (OECD/UNESCO, 2024[29]).
- Chile's Network of Public Innovators connects over 30 000 public servants from all levels of government and other relevant actors for collective learning, creation and experimentation, including on AI.[37]
- Canada's Data Conference serves as the primary forum for public servants and data leaders to enhance awareness, share knowledge and advance data applications throughout the Canadian government.[38] Additionally, department-led working groups on AI topics enable public servants across various departments to share experiences and insights, fostering collaboration and innovation in AI implementation (OECD/UNESCO, 2024[29]).
- The International Smart Cities Network, led by Germany, promotes international exchange and knowledge transfer at national and local level by serving as a place for international dialogue and sharing of ideas and best practices.[39]
- France established ALLiaNCE and *Communauté des labos*; informal inter-ministerial groups for sharing AI best practices.[40]
- In Switzerland, the Competence Network for AI has promoted communities of practice to understand common challenges in the implementation of AI systems, including in public administration.[41]

Such communities and networks need not be specifically focused on AI; in fact, more general groups can help surface a broader base of relevant issues and better consider alternative approaches. However, governments may want to develop additional AI-focused communities and networks or to ensure that general communities and networks include individuals with AI expertise in order to help identify links between problems and AI approaches that may constitute an optimal solution.

Bringing together multi-disciplinary skillsets and perspectives

Digital and AI skills are not the only ones relevant to designing and using AI in government. Some governments have sought to create one or more multidisciplinary teams to ensure AI initiatives benefit from diverse perspectives and expertise. The sensitivity and complexity around AI require the involvement of experts from a range of disciplines, including technology, ethics, law and public policy, to set a strategic approach to the use of AI. Such teams can provide a diversity of perspectives and expertise, thereby facilitating the identification of potential risks, and ensuring comprehensive and inclusive AI use across public administration (Berryhill et al., 2019[4]).

Investing purposefully

Governments are increasingly investing in AI by funding government AI initiatives. Estimates indicate that governments may increase their annual spending on AI-related technologies by 19% in 2025 and continue increasing thereafter (Gartner, 2024[92]). It is essential governments strategically plan, implement and monitor AI investments in government to ensure value for money, identify and mitigate potential investment risks, implement and deploy technologies in a timely manner and evaluate whether intended benefits are realised (OECD, 2025[93]).

The 2023 OECD DGI (2024[74]) shows countries have not yet developed robust capabilities for managing digital investments in the public sector. While 88% of OECD countries have a standardised approach to developing value propositions, only 41% have developed a risk assessment mechanism for digital government investments, including operational (e.g. cybersecurity, service access disruption or related to use of legacy technologies) and financial (e.g. ROI uncertainty, sustainability of funding or overhead and maintenance) risks. To advance towards trustworthy AI investments in government, countries can develop assurance mechanisms, including:

- strengthening strategic planning
- supporting coherent investments across government
- reinforcing investment monitoring mechanisms.

A whole-of government coordination across these three domains will allow governments to invest in AI systems capable of delivering governments' policy objectives on time and on budget. The following sub-sections review these domains.

Strengthening strategic planning for coherent investment

Governments should coordinate with key stakeholders to plan and manage AI investments based on clear principles. Establishing such principles can help to ensure that investment decisions are consistent with overarching strategic objectives. For instance, articulating a commitment to developing more proactive services could result in increased funding and management support for the implementation of AI chatbots for government-citizen interaction. Moreover, coordination among budgeting, digital government and public procurement authorities can help identify AI needs and align them with available resources and potential acquisitions from and partnerships with the private sector. Germany has sought to achieve this through a 2024 AI mission statement that is complemented by a new Centre for AI in Public Administration (BeKI) as an initiating and coordinating body for AI investments and guidance in the federal administration (OECD, 2024[94]).[42]

Governments can also use existing management tools, such as value proposition and investment risks assessment mechanisms to reinforce assurance and secure coherence in investment decisions on AI systems. Adapting value proposition mechanisms to the specificities of AI systems allows governments to strengthen assurance processes for trustworthy development and use of AI. This includes the assessment of key AI aspects such as compliance with regulation and policy standards. The value proposition can include a risk and impact assessment that measures and evaluates the benefits and potential risks of AI systems in government, as well as the plans to comply with regulations. For example, Australia has released a Pilot AI Assurance Framework to guide agencies in aligning AI use cases to Australia's AI Ethics Principles, identify impacts and risks, and apply mitigations.[43]

Funding AI and supporting coherent investments across government

Although often overlooked in national AI strategies, funding and financing mechanisms are an important consideration for government applications of AI (van Noordt, Medaglia and Tangi, 2023[95]). Even simple initiatives need some level of funding and financial support to make their way from idea to reality, with

significantly more funding needed to scale-up a successful project. The availability and nature of this financing can contribute greatly to the eventual success of AI-based innovation. Conversely, a lack of funding for AI development and implementation is a top barrier to government AI adoption (EC, 2024[96]; UK NAO, 2024[16]). Targeted financial resources can support AI experimentation and scaling, as well help in reducing fragmented efforts and uneven adoption of AI. The European Commission (EC) (2024[96]) has recently noted this in a 2024 study on strategic AI adoption for public services, recommending governments increase funding and resources for their use of AI. Examples of specific funding vehicles include:

- The US Technology Modernization Fund (TMF) opened a special call for AI investments to support public agencies as part of a broader funding mechanisms designed to replace outdated legacy IT.[44]

- In France, the Fund for the Transformation of Public Action offers financial support for public sector institutions seeking to improve policies and services with AI, in particular for project proposals with potential for scalability and replicability across the public administration.[45]

- The United Kingdom is investing GBP 110 million to accelerate the AI use in the government, including adding capacity to its Incubator for AI (i.AI), and it has committed to providing GBP 10 million to boost regulators' AI capabilities (Cover-Kus, 2024[97]; UK House of Commons, 2024[98]).

- In Poland, government departments have been asked to set aside a percentage of their budget for AI procurements (van Noordt, Medaglia and Tangi, 2023[95]).

Monitoring mechanisms for coherent investment

Ensuring the on-budget and on-schedule development and deployment of AI investments contributes to the realisation of benefits and delivery of intended results. In line with general investments on digital technologies, governments can leverage monitoring tools to oversee the management and development of AI systems across the administration. These activities should consider developing key performance indicators (KPIs) and structured approaches to manage ongoing developments on AI systems through IT portfolio management. Such management tools can enable or complement guardrails for AI development focused on trustworthy development, deployment and use of AI in government. These approaches can consider quality control mechanisms, linking them with the planning and monitoring of digital investments to secure coherency across the AI system lifecycle. Countries have developed guidance to embed monitoring and measurement actions in the investment cycle of AI initiatives. For example, France uses a monitoring tool to track major digital state projects, including AI initiatives, costing over EUR 9 million.[46] It lists strategic IT projects and helps identify actions for success. The tool monitors project distribution by ministry, progress phase, functional area and estimated cost (OECD/UNESCO, 2024[29]). However, most countries still face the challenge of deploying continuous or ad-hoc monitoring practices. Recognising the need to develop specific capabilities and plans for AI policy monitoring, Norway's Office of the Auditor General (OAG) started auditing the AI use in the central government as part of its pipeline of new performance audits since 2023. At the executive level, the country is taking steps to strengthen the monitoring and oversight of the portfolio of government AI projects through regular internal audits, performance monitoring and impact assessments (OECD, 2024[99]).

Using public procurement to obtain AI products and services and guide the market

Fit-for-purpose public procurement processes and mechanisms are key to enabling agile and cost-effective access to AI systems developed by third parties, ranging from large companies to start-ups and entrepreneurs. Beyond simply purchasing solutions or contracting resources, procurement serves a key strategic approach, which uses purchasing as a bridge to connect public missions and objectives with societal needs and values. To fulfil this role effectively, procurement officials should conduct a comprehensive evaluation of AI's consistency with internal objectives, adherence to fairness and transparency standards, resource efficiency, risk mitigation (e.g. biases or security vulnerabilities),

engagement with stakeholders and impacted groups, and compliance with relevant legal and regulatory frameworks.

In promoting the effective deployment of AI technologies, public entities could consider adopting procurement mechanisms that foster agility, iteration and innovation. The process needs to start with careful preparation and planning to achieve flexible and efficient procurement processes that encourage broad participation which is open and accessible to all (UK DSIT, 2020[27]). This preparatory phase should include:

- the establishment of a multidisciplinary team to support the procurement of the AI systems
- an assessment of current data and governance approaches to evaluate readiness and existing capabilities and resources for the effective training and use of AI systems, as relevant
- an assessment of potential risks throughout AI lifecycle and the identification of associated mitigation strategies.

Agile and innovative procurement methods

Agile and innovative procurement methods provide opportunities to accelerate the adoption of new technologies within governments and on the trustworthy development and use of AI (Monteiro, Hlacs and Boéchat, 2024[100]). These can include technology contests, demonstrations, challenge-based procurement processes and competitive dialogues (UK DSIT, 2020[27]). In addition, policy framework agreements that set overarching procurement rules, priorities and guidelines — either specifically for AI or with key vendors — can play a part in conducting business with the private sector, including for countries with a less diverse set of procurement practices available to them. In Australia, for example, the government used an existing agreement with a multinational technology company to deploy a widescale pilot of an AI solution across its public administration (Australia DTA, 2024[101]). Policy framework agreements can also take the form of predefined environments enabling AI procurement under broader guiding principles. For instance, the EC (2024[96]) Adopt AI programme aims to modernise public procurement for AI systems by fostering dialogue between public procurers and Europe's AI industry. It promotes mutual understanding, drives industry investment and seeks to create a public procurement data space. Sectoral dialogues bridge the gap between procurers seeking solutions and suppliers needing insight into public administration plans (OECD/UNESCO, 2024[29]).

Procurement as a lever for public good and trustworthy AI

Public procurement can be a strategic tool to shape the market and ensure AI systems align with government standards. It also plays a critical role in setting requirements for AI systems that reflect public values, ensuring accountability, security and fairness in AI adoption. For example, the EC set up model contractual clauses to pilot procurements of AI in 2023, which were updated in 2025 to align with requirements of the EU AI Act and provide comprehensive guidelines for high-risk applications and customisable options for non-high-risk AI (2023[102]; 2025[103]).[47] Australia, too, has established model clauses.[48] At the sub-national level, the City of Barcelona, Spain, has introduced procurement clauses emphasising data sovereignty, ensuring that data collected from the public, even by private companies, remains publicly owned (Berryhill et al., 2019[4]). Public procurement guidelines are policy instruments that can influence AI use globally. For instance, ChileCompra, Chile's public procurement agency, has introduced a new tool to ensure procured AI systems are responsible and ethical (see Box 5.24).[49] Internationally, initiatives such as the World Economic Forum's (WEF) (2025[104]) "AI Procurement in a Box" provide structured guidance to help governments integrate AI procurement best practices and align AI acquisitions with ethical and regulatory frameworks. In April 2025, the United States issued new AI acquisitions policy that seems to promote agile procurement and the removal of unnecessary bureaucracy and outdated procurement processes (Box 4.7).

Box 4.7. Driving efficient AI acquisitions in the United States

In the United States, the White House Office of Management and Budget (OMB) issued *M-25-22 Driving Efficient Acquisition of Artificial Intelligence in Government* on 3 April 2025. It includes a variety of requirements and recommendations for federal agencies regarding AI acquisitions across six procurement phases. A non-exhaustive list of these include:

1. **Identification of requirements**: convene a cross-functional team to inform the procurement of AI systems and assist in creating an initial list of potential risks to be evaluated. As practicable, consider which uses may be "high-impact AI" (Box 1.3).

2. **Market research and planning**: seek state-of-the-art AI capabilities by conducting thorough market research, including through interagency knowledge sharing and considering novel capabilities from new entrants. Seek detailed demonstrations and tests of potentially useful AI to assess providers and identify obstacles to long-term cost effectiveness. Use performance-based techniques to identify requirements and contract terms to understand and assess vendor claims.

3. **Solicitation development**: include in solicitations requirements that protect against vendor lock-in and terms related to IP rights and lawful use of government data, and when practicable, agencies must be transparent regarding whether the AI use could be considered "high-impact" and what this could mean to the vendor.

4. **Selection and award**: test proposed solutions to understand their capabilities and limitations. Separately, agencies must evaluate proposals to identify any potential new AI-related risks that were not previously identified. Address in contract terms, where applicable, IP rights and the use of government data, privacy, vendor lock-in protections, compliance requirements for the policy discussed in Box 4.1, ongoing testing and monitoring, vendor performance requirements

5. **Contract administration**: help ensure AI systems are authorised by an appropriate official prior to deployment, put in place contract oversight and monitoring processes for contract performance and to identify and mitigate emerging risks. Arrange for periodic evaluation of the AI system or service's value to the government, considering, as practicable, effectiveness, efficiency, risks, operations and maintenance costs and stakeholder feedback. Consider sunset criteria.

6. **Contract close-out**: help ensure vendor lock-in protection, such as ensuring ongoing rights and access to any data or derived products.

To assist agencies, a centre of government agency will release publicly available guide(s) to assist the acquisitions workforce with the procurement of AI systems, create a digital repository available to public servants to facilitate the sharing of information, knowledge and resources about AI acquisitions (e.g. best practices, tools, language for contract clauses, negotiated costs).

Source: (US OMB, 2025[105]).

Procurement to shape the market

Beyond ensuring governments procure trustworthy AI for their own use, public procurement can act as a powerful lever to influence more general market dynamics, driving innovation and aligning AI system development with principles for trustworthy AI. Public procurement represents about 13% of GDP in OECD countries (OECD, 2024[106]). By approaching procurement strategically, governments can use the "economic weight of the government's purchasing power" to foster the development of AI solutions that

not only meet their needs but also promote broader alignment with ethical and regulatory standards (World Bank, 2025[107]).

Expanding AI's potential through partnerships

Governments can benefit greatly from ongoing, active cross-sector partnerships in which each sector has a concrete role and contributions (OECD/CAF, 2022[30]). These partnerships can facilitate collaboration among public entities and AI specialists in other sectors, including private sector companies and non-governmental actors (e.g. academic institutions, foundations), promoting the development and implementation of cutting-edge solutions. Public-private partnerships (PPPs) are perhaps the most common type of arrangement. Some examples here include:

- The European Union's InvestAI initiative seeks to mobilise EUR 200 billion in investment in AI through a PPP akin to a CERN for AI to enable to development of leading-edge AI systems across sectors.[50]

- As announced at France's February 2025 AI Action Summit, 10 countries are developing a Public Interest AI Platform and Incubator to support, amplify, decrease fragmentation between existing public and private initiatives on public interest AI, and address digital divides. It will support digital public goods, technical assistance and capacity building to foster a trustworthy AI ecosystem for public interest AI.[51]

- Chile's Data Observatory (DO) is a non-profit organisation jointly led by government, industry and academia. It serves as is a technological centre that, through the management of large volumes of data, seeks to contribute to social well-being by promoting sustainable development of the country, contributing to the generation of enabling factors for optimal development of AI and promoting the creation of public policies and strategic decision-making based on evidence.[52]

- Portugal's ChatGPT-driven virtual assistant for public services (see Box 5.46) was developed through a PPP with the government and several companies.

- Latvia's new Artificial Intelligence Centre is a private foundation co-founded by government, academia and industry, designed to promote the trustworthy and sustainable adoption of AI across sectors, including a specific focus on the integration of AI in public administration, and to ensure the incorporation of Latvian language and culture in AI systems.[53]

Turning to GovTech startups

Bridging the concepts of public procurement and partnerships, GovTech is the collaboration between government and start-ups, innovators, government "intrapreneurs" and academia on innovative digital government solutions. It complements existing government capability for agile, user-centric, responsive and cost-effective processes and services (OECD, 2024[108]). It aims to contribute to an agile government and enhance digital government maturity. Not only does this help improve effectiveness and efficiency, but it can also encourage the participation of start-ups and newer providers in the government market. GovTech innovation is characterised by co-creation and experimentation. These collaborative interactions aim to transcend traditional supplier-contractor relationships to build new forms of partnerships. Rather than focusing on the detailed terms of reference and technical specifications, GovTech's focus is instead on the solution's expected outcomes and on involving GovTech actors in building it. While many such collaborations leverage public procurement (as discussed above), they can also use grants and monetary prizes to incentivise the creation of innovative solutions (e.g. through demo days or incubator programmes). The OECD has developed a GovTech Policy Framework to outline the factors important for maximising GovTech engagements (Figure 4.4).

Figure 4.4. OECD GovTech Policy Framework
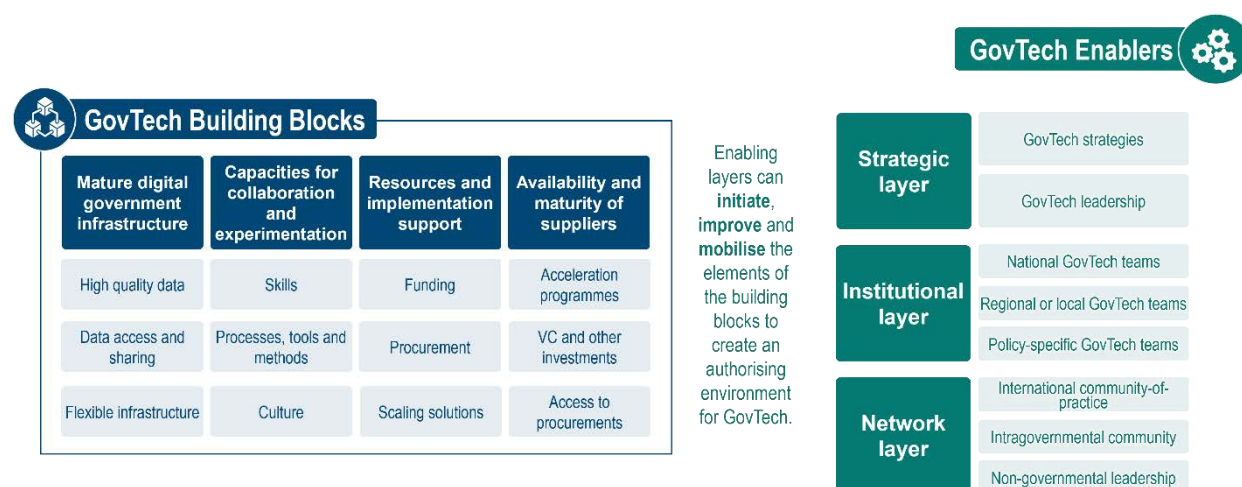
Governments can leverage GovTech collaborations to experiment with and develop AI systems to address governmental and societal challenges. The 2023 OECD Digital Government Index (DGI) shows that 42% of the 33 OECD countries surveyed are setting GovTech objectives to facilitate the testing and adoption of emerging technologies, including AI. For example, Spain (2024[109]) is using its GobTech Lab to develop AI pilots. A recent EC report (2024[110]) titled *GovTech: influencing factors, common requirements and recommendations* provides several other use cases and findings.

Establishing guardrails to guide strategic and responsible AI

Guardrails help to ensure the trustworthy deployment, development and use of AI in government. They can be **binding and non-binding policy levers, transparency processes and accountability mechanisms, such as monitoring and oversight bodies**. Guardrails are essential for managing the risks associated with AI and deploying AI according to legal boundaries and social values. This ultimately helps to build public trust in government. The sections below review key guardrails, together with available policy options that governments can consider implementing in their own contexts, drawing on examples of international good practices.

It is important to note, however, that these guardrails should be seen in conjunction with the enablers above. Guardrails and their interpretation can be a leading driver in risk aversion, which contributes to risks of inaction and missed opportunities (see Chapter 1). Governments should also consider the importance of eliminating or revising guardrails that no longer serve their purpose or cause negative consequences that do not outweigh their utility.

Finally, this section does not seek to suggest that governments should necessarily put in place all guardrails discussed in this section, nor that they should apply to all uses of AI. That, too, could contribute to risk aversion and hinder the trustworthy adoption of AI in government. Instead, governments should determine which guardrails fit their operations and contexts and apply them to AI uses in a risk-based way commensurate and proportionate to their potential level of risk.

Using policy levers to guide trustworthy AI

Policy levers to promote trustworthy AI can include **non-binding instruments** such as guidance documents, ethical frameworks, technical standards and risk-management frameworks; and **binding instruments**, such as laws and regulations. These policy levers work together and alongside other guardrails to protect human rights. They do so through risk management approaches and by fostering the responsible development and deployment of AI. They help mitigate AI misuse, skewed outcomes, privacy infringements and unintended consequences, while also offering practical tools for implementing governance principles and ensuring consistent performance across AI use (UNESCO, 2024[111]). Generally, developing policy levers for AI in government should follow a set of good practices to ensure they effectively promote trustworthy AI. The policy levers should:

- Align with ethical principles and societal values to ensure that technology serves the common good. This is important for maintaining public trust, safeguarding the free exercise of human rights and ensuring that AI systems operate fairly and responsibly.

- Take into account both innovation and risk management, helping government organisations navigate the evolving AI landscape responsibly by harnessing AI's transformative potential while addressing risks and challenges (OECD, 2023[35]; 2021[112]).

- Continuously assess and identify potential risks associated with AI systems. Risk assessment tools offer a balanced approach, where both responsible use and cutting-edge advancements can coexist, ensuring that AI benefits society while meeting ethical standards (UNESCO, 2023[113]).

- Engage stakeholders through different institutional arrangements to align AI use with the needs and values of those it impacts (see Engagement section of this chapter).

Non-binding policy levers

One of the most common entry points for promoting trustworthy AI is the adoption and/or development of principles or ethical frameworks. These instruments establish a set of values and best practices to guide AI that is used transparently, fairly and responsibly. Around the world, over 200 such instruments have been developed (Corrêa et al., 2023[114]), and often times, such principles are embedded in a country's national AI strategy. They also often address a wide range of ethical concerns surrounding AI, such as bias, transparency, accountability and the impact of AI on society.

Important progress has been made in global governance for AI, with several instruments developed by intergovernmental and supranational organisations to standardise and unify AI development on a global scale. These include the OECD AI Principles (see Table 2.2 in Chapter 2), the EC's Ethics Guidelines for Trustworthy AI, and the G7 Hiroshima AI Process International Guiding Principles for Advanced AI Systems.[54] The African Union (AU) is also working on its own charter on trustworthy AI (OECD.AI, 2025[115]). Other non-binding international efforts include the UN General Assembly's 2024 resolution on the promotion of "safe, secure and trustworthy," which was backed by more than 120 countries,[55] as well as declarations from international AI summits, which have been held in the UK, Korea and France.[56] At the national level, governments also establishing their own ethical frameworks. For example:

- Australia has the "Artificial Intelligence Ethics Principles", which are designed to prompt organisations to consider the impact of using AI enabled systems and help businesses and governments to practice the highest ethical standards when designing, developing and implementing AI. Building upon this framework, Australia's Voluntary AI Safety Standard (VAISS) gives practical guidance to all Australian organisations on how to safely and responsibly use and innovate with AI.[57]

- In Colombia, the "Ethical Framework for Artificial Intelligence" offers a set of principles and a methodology that should be considered in the design, development and deployment of AI systems (OECD, 2024[3]).[58]

- Egypt has developed an "Egyptian Charter for Responsible AI" shaped around key principles.[59]

Governments are increasingly developing guidance documents for using AI. These are comparable to the guidance discussed above that serve as enablers, but with more of a focus on establishing the parameters for trustworthy use of AI. Usually geared towards public officials responsible for developing AI projects and managing extensive data collection and analysis, these documents are intended to equip public officials with knowledge needed to ethically shape AI projects and to raise awareness about potential risks, including, but not limited to, breaches of personal data (OECD/UNESCO, 2024[29]). Such guidance is more concrete than principles, often addressing technical aspects that may affect AI deployment. For example:

- In the UK, the government and The Alan Turing Institute jointly developed a guide to "Understanding AI Ethics and Safety" (2019[116]). In Canada, the "Guide on the Use of Generative AI" serves as a resource for federal institutions utilising generative AI technologies (OECD/UNESCO, 2024[29]).[60]

- Germany has two primary sets of guidelines to ensure the ethical use of AI in public services: Guidelines for the Use of AI in Employment and Social Protection Services; and the AI Guidelines for Federal Administration (OECD/UNESCO, 2024[29]; Policy Lab Digital, Work & Society within the German Federal Ministry of Labour and Social Affairs, 2024[117]). The latter has already been published, while the former is under development.[61]

It is important to note that non-binding measures are limited in what they can achieve. For instance, when it comes to AI in the workplace, most OECD countries' AI-specific measures to promote trustworthy AI in the workplace are primarily non-binding and rely on organisations' capacity to self-regulate (i.e. soft law) (OECD, 2023[118]). Because of its non-binding nature, soft law may not be enough to prevent or remedy AI-related harm in the workplace. Government should also consider binding measures to overcome this limitation in important areas.

Binding policy levers

To date, most binding measures for AI in government have been put in place at national and sub-national levels, as discussed below. However, some **international mechanisms** have recently come into action. Perhaps the most notable example is the EU AI Act regulation (Box 1.2). More recently, the Council of Europe (CoE) Framework Convention on AI and Human Rights, Democracy and the Rule of Law (2024[119]) was passed as the first international legally binding treaty on AI. Opened for signature in September 2024, it applies to both the public and private sectors. It aims to ensure that activities within the lifecycle of AI systems are fully consistent with human rights, democracy and the rule of law, while being conducive to technological progress and innovation. As of September 2025, the treaty has 17 signatories, including the European Union and the United Kingdom, as well as non-European countries, such as Canada, Japan and the United States.

National AI laws and regulations can govern activities throughout the AI lifecycle and address issues like data protection, privacy, misuse and other concerns. Such rules help ensure that AI development aligns with societal values and legal norms by providing developers with clear guidance for compliance, establishing boundaries through binding standards for transparency, accountability and fairness (OECD, 2025[120]). They can also define responsibility for AI outcomes and promote consistency and cooperation across jurisdictions by harmonising standards. Ultimately, adequate and well-fitting laws and regulations help promote innovation while instilling necessary protections, ensuring that AI serves the public interest. These binding levers can be general AI governance laws that affect all sectors, AI-specific laws focused on government use, and cross-cutting laws related to but not specifically targeting AI.

However, governments need to keep in mind the dynamic nature of AI in the present and its many potential trajectories in the future. As discussed in Chapter 3, public servants face challenges with confusing or outdated rules that hinder their ability to adopt AI. In developing national binding instruments, they should

seek alignment with the OECD (2021[121]) *Recommendation Agile Regulatory Governance to Harness Innovation*. The United Kingdom has sought to achieve this through its "pro-innovation approach to AI regulation".[62]

General AI governance laws are broad rules that govern AI systems across various economic sectors, policy domains and regions. They focus on establishing frameworks for AI risk management, ethical use and societal impact. While many countries have laws that may influence AI (e.g. data protection laws), few have formal laws specifically on AI. Existing or proposed national laws and regulations include:

- Korea's Basic Act on the Development of AI and the Establishment of Foundation for Trustworthiness, or "AI Basic Act" (2024[122]), which will come into effect in January 2026, establishes a comprehensive framework to promote AI innovation while ensuring ethical standards, safety and public trust for all organisations using AI in the Korean market.
- Bahrain and Oman put forward draft AI legislation in 2024, with Oman holding a public consultation.[63]
- Governments in Latin America (including Argentina, Brazil, Chile, Colombia, Costa Rica, Ecuador, Mexico, Panama, Peru and Uruguay) are discussing general AI legislation (UNESCO, 2024[111]).

Specific laws and regulations can also be developed to govern the use of AI systems in government. These often focus on transparency, accountability, AI governance in public organisations, ethical considerations and responsible AI use in functions and services. In the United States, at the sub-national level, states have taken steps in this direction. For instance, New York introduced the LOADinG Act in 2024 to limit the use of automated decision-making systems by state agencies and provide some protections for public servants as related to AI (Werner, 2024[123]). Delaware has strengthened oversight of AI and generative AI in the state (2024[124]). However, the adoption of dedicated AI laws or regulations for government use remains limited, with most frameworks emerging as part of broader AI governance efforts. Besides hard laws and regulations, formal policy guidance can also provide binding rules for government organisations. For example, the 2025 US policy on "Accelerating Federal Use of AI through Innovation, Governance and Public Trust" covers a range of issues and measures to ensure trustworthy AI in government while removing barriers to innovation (Box 4.1).[64]

Cross-cutting laws are broader laws that, while not exclusively for AI, have significant implications for its deployment and use in government. These might include regulations on data protection, privacy, cyber risks and human rights, which shape how AI systems can be used in public sector contexts. For instance, data protection and privacy law safeguard personal information by setting rules on how AI systems collect, store and process data, ensuring individuals' privacy rights are upheld. While not typically designed specifically for AI, they have significant implications for its use. For example, the EU GDPR outlines specific requirements for managing personal data for all sectors, covering aspects such as data collection, storage, processing and the rights of individuals.

Promoting transparency in how government uses AI

To be transparent about AI use, governments should, to the extent practicable and appropriate, make algorithms open, understandable and accessible to public scrutiny, as well as disclose the processes and decisions AI systems contribute to. This means governments should provide clear, context-appropriate information about how its AI systems work, the data they use and how they reach conclusions and outputs, and mechanisms to challenge outcomes. Such transparency allows stakeholders to access information, make informed decisions, and if necessary, seek redress for potential harms. Designing a coherent, sustainable and impactful approach on AI algorithmic transparency involves disclosing key information about AI systems and algorithms used. This includes complementary assets such as training data; engaging with a broad and diverse set of stakeholders to ensure their needs and concerns are considered (see the "Engagement" section of this chapter); promoting AI literacy to empower communities to have an

informed voice in the issues that affect them; and strengthening existing national rules on transparency and accountability to effectively address the challenges and risks posed by AI.

Transparency not only promotes trust and enhances public value but also underpins government accountability. As covered by the OECD AI Principles, transparency and accountability are two different but complementary concepts. Transparency enables more informed oversight, fosters trust and increases the accountability of those who develop or control these systems. Transparency also enhances the fairness of AI-driven systems and helps ensure that their implementation can be effectively monitored and evaluated, particularly when decisions have a direct impact on people's lives in areas such as healthcare, finance and criminal justice.

Countries should employ a variety of transparency policies, tools, methods and approaches that are suited to their audience and provide clear information. These instruments can be broadly classified as **either proactive or reactive** (GPAI, 2024[125]).

Proactive transparency instruments

Overall, "the public sector's understanding of its own AI usage is severely lacking, which hinders both democratic accountability and internal knowledge sharing" (Ada Lovelace Institute, 2025[23]). Instruments are needed that allow governments to understand its own AI use, and by extension, enable them to proactively share information about the AI systems used in the public administration without being prompted by requests (GPAI, 2024[125]). Policy options for proactive transparency include public registries of AI systems, publishing algorithm source code and documentation, user-driven proactive publications and automated responses triggered by interactions.

Public registries of AI systems are increasingly common, serving as centralised repositories that consolidate information about AI systems currently used in government. The goal is to create a "one-stop shop" where citizens and stakeholders can easily access information about the AI systems in use, their purposes, the sectors they apply to and the jurisdictions they affect. Examples include:

- Colombia's dataset on automated decision systems in the Colombian public administration[65]
- the United Kingdom's Algorithmic Transparency Records[66]
- the US government's AI use case inventory, which federal agencies are required to update at least annually (US OMB, 2025[17])[67]
- national government public algorithm inventories in Chile, France and the Netherlands Public Algorithms Inventory[68]
- sub-national algorithm registers Amsterdam, the Netherlands, and Helsinki, Finland.[69]

Developing a central, public and searchable registry of AI systems is a best practice that enhances transparency. Yet doing so can be challenging, in part due to how rapidly governments are deploying AI systems in a variety of domains. In Chile, challenges in making AI uses cases transparent prompted the Chilean Transparency Council, an independent body created by law, to issue recommendations on improving algorithmic transparency in government (2024[126]). In the Netherlands, only about 5% of AI systems have been published in the Dutch AI registry, as of October 2024 (Netherlands Court of Audit[127]). In the United Kingdom, the Public Accounts Committee (PAC) (2025[128]) found that relatively few (33) Algorithmic Transparency Records had been published, jeopardising public trust in the adoption of AI in government. However, such registries could be populated automatically, depending on if and how impact and risk assessments are carried out in a government (OECD, 2024[94]). For instance, Canada's Directive on Automated Decision-Making requires the completion of an Algorithmic Impact Assessment (AIA) for automated decision systems. The AIA results must be published as open information on Canada's Open Government Portal. If required for all AI systems, such a process could automatically populate a public registry.

Publishing algorithm source code and documentation also promotes transparency. Open sourcing the code for public algorithms is regarded as a best practice in algorithmic transparency and is especially valuable for technical and expert audiences (Ada Lovelace Institute, 2021[129]). This allows those with the necessary skills to examine, test and verify how these systems operate, promoting accountability and trust. Some efforts are a step shy from releasing full source code but require the publication of thorough documentation than can have a similar effect.

- In France, the Digital Republic law mandates government agencies to "make publicly available, in an open and easily re-usable format, the rules defining the main algorithmic processing used in the accomplishment of their mission when such processing is the basis of individual decisions".[70]
- The United Kingdom's the Algorithmic Transparency Recording Standard (ATRS) mandates public sector organisations to transparently disclose details about their use of algorithmic methods in decision-making processes (OECD, 2023[130]).[71]
- In Canada, the Directive on Automated Decision-making explicitly outlines, in detail, explainability requirements for AI systems differentiated by levels of risk, determined by use of an Algorithmic Impact Assessment tool (the results of which must also be published).[72]
- The US policy discussed in Box 4.1 requires federal government agencies, when practicable and subject to some exclusions, to release and maintain AI code as open-source software in a public repository.

Transparency efforts can also be iterative, as seen in **user-driven proactive publications**. This type of proactive publications involves public entities choosing to proactively disclose information after receiving numerous similar requests. By publishing this information proactively, future requests are avoided, saving time for both officials and requestors. While its use in ensuring algorithmic transparency is not well-documented, this approach could be a relevant and cost-effective method for disclosing frequently requested information related to algorithms (GPAI, 2024[125])

Some types of disclosures may only be made for some users in context-specific situations. **Automated responses triggered by interactions** occur when information about an automated decision-making system is automatically provided during specific governmental processes. For instance, when someone engages with a public body's website or online platform for a service or administrative procedure involving an automated decision-making system, relevant information about the system could be automatically disclosed, without the user needing to request it explicitly (GPAI, 2024[125]).

Reactive transparency instruments

Reactive disclosure transparency instruments allow government to respond to specific requests for information from individuals, groups or authorities. Unlike proactive disclosure, this approach is initiated by external demand rather than the government's proactive disclosure (GPAI, 2024[125]). These involve submitting a request under the country's relevant Access to Information (ATI) law to obtain information about an algorithm or its use, leveraging an existing policy instrument widely available in most contexts or countries.[73] However, these regimes are not specifically designed for algorithmic transparency and can be ineffective when applied in this context (Valderrama, Hermosilla and Garrido, 2023[131]). For example, requesting information about an algorithm's source code or its application is unlikely to produce the desired results due to issues with record management practices and common exceptions in ATI laws, such as conflicts with intellectual property restrictions and trade secrets of private providers of public services (Fink, 2017[132]; Brauneis and Goodman, 2017[133]).

Advancing accountability through risk management throughout the AI system lifecycle

For some government AI systems, the context of their development or use may pose a higher risk. This can relate to their scale (seriousness and probability of adverse impact), scope (breadth of application,

such as number of individuals affected) or optionality (degree of choice as to whether to be subject to the effects of an AI system) (OECD, 2022[58]). Risk management procedures can help to identify which systems or contexts pose higher risks to mitigate them (OECD, 2023[134]).

Risk management for AI systems that may carry high risks should be informed by guidance on which levels of risk are acceptable for different uses and contexts. Risk management is needed both before — such as through ex ante impact and risk assessments — and after the deployment of AI systems. One of the most well-known examples is the US National Institute of Standards and Technology (NIST) (2023[135]) AI Risk Management Framework. This framework helps public or private organisations identify unique risks posed by generative AI and proposes actions for generative AI risk management that best align with their goals and priorities. While designed for the US, it has been translated in several other languages and used in other countries. The G7 Hiroshima AI Process International Guiding Principles for Advanced AI Systems and Code of Conduct also sets baseline standards to manage risks (2023[136]; [137]). The US also requires risk assessments for AI use and puts in place risk management practices for uses deemed "high-impact" (see Box 1.3) (US OMB, 2025[17]). As another national example, Türkiye's Digital Transformation Office conducts the "AI Risk Management Recommendation" and "Trustworthy AI Seal" studies to closely monitor the use of AI for public benefit (OECD, 2024[3]).

AI experts recommend that governments make establishing or adopting such processes a top priority for mitigating AI harm (OECD, 2024[138]). Yet the proliferation of frameworks can make it difficult for governments to determine which is the most appropriate to follow. As calls for the development of risk management frameworks continue to grow, interoperability would enhance efficiency and reduce enforcement and compliance costs. The OECD (2023[139]) is actively working to promote policy coherence and interoperability among these frameworks.[74]

Impact assessments

Impact assessments, including Algorithmic Impact Assessments (AIA), can help public organisations to anticipate and evaluate how an algorithm may function in a specific context. They are evaluations of an AI system that use prompts, workshops, documents and discussions with the developers of an AI system and other stakeholders to explore how the system will affect people or society in positive or negative ways (Valderrama, Hermosilla and Garrido, 2023[131]). These tend to occur in the early stages of a system's development before it is in use (ex-ante) but may occur after a system has been deployed (ex-post).

The primary aim of ex-ante AIAs is to assess the potential impacts of an algorithmic system on economies and societies and to provide a mechanism for accountability (Valderrama, Hermosilla and Garrido, 2023[131]). AIAs also help to better understand, classify and mitigate potential risks or harms associated with the algorithm. An example employing such a technique is Canada's "Directive on Automated Decision-Making," which requires an AIA that considers various factors and, in turn, provides a risk score that prescribes certain actions. The approach has been adapted in other countries, like Uruguay, where it informed their Guide for Algorithmic Impact Study (OECD/CAF, 2022[30]). In 2024, the Council of Europe issued the Human Rights, Democracy and Rule of Law Impact Assessment (HUDERIA).[75] Its methodology provides for the creation of a risk assessment and a mitigation plan to minimise or eliminate the identified risks, protecting the public from potential harm.

Ex-ante AIAs are the most common approach in use today. The OECD has generally been supportive of this approach because they help to convert principles into actions.[76] However, some argue that evaluating impacts is not the same as evaluating harms, and that in some instances, doing so can obscure true harms. That is, in part, because the metrics used for impact assessments often are not measuring emotional or psychological harm (Gupta et al., 2021[140]). This work also suggests that AIAs often do not consider the voices of all who may be impacted by an AI system. Other critics of AIAs argue that these assessments are not designed to continuously monitor the effects of deployed systems and adapt accordingly based on feedback and real-world ramifications (Mehta, Rogers and Gilbert, 2023[141]). Recent findings from the Ada

Lovelace Institute (2025[23]), on lessons learned from six years of studying AI in government, further underscore this point. One of its main findings is that "AI is 'sociotechnical', in that it influences and is influenced by the social contexts in which it is deployed, often with unintended ripple effects. The success and acceptance of AI tools depends on their interaction with existing social systems, values and trust. Focusing exclusively on technical criteria while failing to consider these factors can lead to scepticism, and ultimately hinder adoption and use".

To ensure AIAs are valuable, governments will need to conduct thorough sociotechnical assessments that engage appropriate stakeholders and integrate diverse perspectives (Lam et al., 2023[142]). However, ex-ante evaluations alone are often insufficient; they should be complemented by ex-post impact assessments that build on top of the ex-ante AIAs. This implies the development of mechanisms for continuous monitoring, adaptation and accountability, ensuring that AI systems evolve in response to real-world evidence. The US NIST has launched an Assessing Risks and Impacts of AI (ARIA) to advance sociotechnical testing and evaluation for AI.[77]

As related to ex-ante impact assessments, governments should consider in advance whether AI is the best solution to address a given problem, as discussed in the enablers section on "Determining whether AI is the best solution".

Algorithmic audits

After deployment, it is important that governments continue monitoring system behaviours to determine what expected or unexpected risks may be materialising, and to ensure that government organisations carry on with implementation in a responsible manner. This usually takes the form of algorithmic audits. These involve independent, usually external, scrutiny of an AI system or the processes around it. These can be technical audits of the system's inputs or outputs; compliance audits of whether an AI development team has completed processes or regulatory requirements; regulatory inspections by regulators to monitor behaviour of an AI system over time; or sociotechnical audits that evaluate the ways in which a system is impacting wider societal processes and contexts in which it is operating. Because audits usually occur after a system is in use, they serve as accountability mechanisms to verify whether a system behaves as developers intend or claim. Examples of algorithmic auditing process can be seen in Box 4.8.

Governments need to design their audits carefully, however. Inadequacies in AI auditing could create false confidence and "obscure problems with algorithmic systems and create a permission structure around poorly designed or implemented AI" (Goodman and Trehu, 2022[143]). Some experts argue that insufficient audits may prove meaningless or could exacerbate the problems they are designed to address, as well as being used as "audit washing" to give the appearance of due diligence.

AI system capabilities risk assessments

AI systems capabilities risk assessments are similar to impact assessments but look specifically at the likelihood of harmful outcomes occurring from an AI system due to a system's capabilities. These also tend to occur in the early stages of a system's development before it is in use but may occur after a system has been deployed. Such approaches should take into account risks related to AI systems' limitations and capabilities, as well as contexts of use. For example, the government of Queensland, Australia issued a "foundational AI risk assessment guideline" for public servants.[78] Risk assessment have become increasingly common in the private sector for advanced AI systems, such as with "responsible scaling policies" (RSPs), which commit to actions based on risk assessment of AI system capabilities (OECD, 2024[138]). When identifying potentially dangerous capabilities, RSPs often set thresholds that trigger actions to slow or cease development (METR, 2023[144]).[79]

Testing and assessment bodies and national, multilateral or regional bodies, such as AI safety or security institutes, are increasingly playing a role in facilitating risk management, including through building testing

and assessment ecosystems (OECD, 2024[138]). For instance, the UK DSIT provided guidance on responsible capability scaling.[80]

Relative to the abundance of examples on governments using impact assessments or algorithmic audits, there are few examples of governments developing or using other AI risk assessments for AI in government. This may be because most government AI systems are procured from private sector companies, which may conduct risk assessments prior to putting their products on the market. It may also be because government AI systems, as discussed in Chapter 2, often do not leverage the latest approaches and instead rely on rules-based systems or more established ML approaches. Still, governments should consider and use such risk assessment approaches as they seek to use more advanced and capable AI systems. For example, in addition to structured testing, governments could conduct adversarial testing—commonly known as red-teaming—especially for complex foundation models used or acquired by governments, to proactively identify vulnerabilities, misuse risks and harmful system activities or outputs in sensitive government applications. Capability assessments could go also include language-specific evaluations, particularly for LLMs used in multilingual public services. These assessments could help ensure fair performance, address data imbalances in foundation models and help verify that models are suitable for the languages they serve.

Empowering oversight and advisory bodies to guide responsible AI

Oversight entities

The role of oversight bodies is evolving to meet new challenges posed by the expansion of AI use across government. For instance, Supreme Audit Institutions (SAIs)[81] are increasingly required to extend their audit activities to include the scrutiny of AI algorithms that underpin government operations and public decisions. This shift necessitates a comprehensive evaluation of AI algorithms not only for their accuracy, security and effectiveness but also for their transparency and fairness. Box 4.8 illustrates how SAIs have adapted their role to conducting algorithmic audits and developing the frameworks for doing so.

Audits serve a range of purposes, including evaluating the performance of algorithmic systems against established standards, ensuring regulatory compliance, detecting unlawful discrimination, enhancing transparency and explainability, assessing security and robustness, evaluating broader social and ethical impacts and holding organisations accountable for their systems. Additionally, audits may be used to identify systemic failures in the use of algorithmic systems, offering valuable insights that can inform their application in another context (Ada Lovelace Institute, 2021[129]).

Box 4.8. Public sector algorithmic auditing approaches and tools

France

In 2024, the French *Cour des Comptes* evaluated the integration of AI within the Ministry of Economy and Finance. Since 2015, the ministry has implemented 35 AI programmes aimed at detecting individual fraud risks, identifying business difficulties and providing faster responses to users. While technological aspects are well managed, the report found that ethical, human resources and environmental considerations remain underexplored. The Cour recommended robust ministerial oversight to ensure trustworthy public AI, better assessment and transparent allocation of productivity gains, and proactive anticipation of AI's impact on staff roles

Netherlands

The independent Netherlands Court of Audit has audited the Dutch government's use of algorithms. Through its evaluation, the Court developed an audit framework specifically designed to assess the use

---

of algorithms within government. The framework evaluates a wide range of aspects, from governance and accountability to technical aspects like the AI systems and data, privacy, IT general controls and ethical considerations. The framework is being used across several Dutch institutions to guide their development of new algorithms. In 2022, it audited nine major public sector algorithms and found that six (67%) did not meet basic requirements, exposing the government to bias, data leaks and unauthorised access.

Sweden

The National Audit Office of Sweden conducted an audit of three automated decision-making systems used by the Swedish Government: the parental benefit at the Swedish Social Insurance Agency, annual income taxation of private individuals at the Swedish Tax Agency and driving licence learner's permits at the Swedish Transport Agency. The audit aimed to assess whether these systems operated effectively and efficiently while safeguarding legal certainty in decision-making. It evaluated the performance of the systems against legislative standards for efficiency and legal certainty, identifying specific shortcomings.

United Kingdom

In 2024, the UK National Audit Office (NAO) published a Report on AI in Government examining how effectively the UK government bodies are using AI for public services. It found that only 21% of the 87 bodies analysed had an AI strategy, as required by policy, while 61% had plans to develop one. Notable initiatives include the Department for Work and Pensions' establishment of an AI steering board and a separate advice and assurance group, and the Ministry of Justice setting up of an AI steering group to review individual AI use cases, coupled with the adoption of algorithm consultation panels, including end users and data ethicists.

United States

The US Government Accountability Office (GAO) issued *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*, which identifies key accountability practices — centred around the principles of governance, data, performance and monitoring — to help federal agencies and others use AI responsibly.

Cross-border collaboration

The Supreme Audit Institutions (SAIs) of Finland, Germany, the Netherlands, Norway and the United Kingdom collectively issue *Auditing machine learning algorithms: A white paper for public auditors*, which is updated over time.

Source: https://www.ccomptes.fr/fr/publications/lintelligence-artificielle-dans-les-politiques-publiques-lexemple-du-ministere-de, https://english.rekenkamer.nl/publications/reports/2021/01/26/understanding-algorithms, https://www.riksrevisionen.se/download/18.2008b69c18bd0f6ed3f25040/1608291082190/RiR_2020_22_en-GB.pdf, (Ada Lovelace Institute, 2021[129]), https://www.gao.gov/products/gao-21-519sp, https://www.auditingalgorithms.net, (OECD, 2023[130]), https://www.nao.org.uk/reports/use-of-artificial-intelligence-in-government.

Oversight bodies serve as platforms that bring together diverse expertise and perspectives, which is important for the effectiveness of any accountability mechanism (Ada Lovelace Institute, 2021[129]). For example:

- In 2023, Spain created the independent Spanish Agency for the Supervision of AI (AESIA) (Pehlivan and Valín, 2023[145]).
- In 2024, the EU AI Act established a supervisory European Artificial Intelligence Board (AI Board).[82]

- Ombudsman are playing a key oversight role in public organisations' use of AI. For instance, the European Ombudsman has investigated the use of AI by the EC,[83] and the Dutch Ombudsman and Data Protection Authority have actively examined algorithmic decision-making and AI's impact on citizens' rights in the Netherlands.[84]
- Parliaments and their oversight committees are taking a more active role in different countries, including Australia and the United Kingdom,[85] where they have conducted inquiries into AI and algorithmic transparency. Some parliamentary bodies establish ad hoc committees to investigate specific AI-related concerns, particularly when emerging risks or controversies arise.

Advisory bodies

Advisory bodies can also help to ensure governments are using AI in a trustworthy manner. They can provide expert guidance, insights and recommendations in response to specific requests from government on emerging issues in AI. Some can be more hands-on, with Ireland's AI Advisory Council, for example, developing and delivering its own workplan of advice.[86] Other examples include:

- New Zealand Data Ethics Advisory Group offers non-binding guidance on the use of algorithmic systems by public agencies. Its recommendations address issues such as human rights compliance, scientific validity, privacy and ethics (Ada Lovelace Institute, 2021[129]). [87]
- Greece established a High-Level Advisory Committee on AI in November 2023, which plays a pivotal role in shaping national AI policy. It focuses on promoting economic and societal growth while addressing the risks associated with unchecked AI use. It has developed "A Blueprint for Greece's AI Transformation" (2024[47]). The country has also established a National Committee of Technoethics and Bioethics to provide independent expertise in providing strategic guidance and recommendations on the ethical implications of AI, among other things.[88]
- Spain (2024[146]) has created the Artificial Intelligence Advisory Council as a formal independent body to provide the government with analysis, advice and support on the topic of AI. It held its first meeting in June 2024.
- The government of Western Australia formed an AI Advisory Board in 2025 to provide advice to Western Australian government agencies on risk mitigation and to support the safe, responsible and ethical use of AI in the Western Australian public sector.[89]

AI safety and security institutes and units

Governments around the world are also focusing on this topic through efforts including the establishment of AI safety and security institutes and units in several countries (OECD, 2024[138]). For instance, Canada, the EU, France, Japan, Korea Singapore, the UK and the US have each launched such an institute or unit,[90] with these and three additional countries deciding to form an international network of institutes (UK DSIT, 2024[147]). The mandate of such institutes or units is generally broader than AI's use in government, but some do include a focus on this. For instance, the upcoming IndiaAI Safety Institute will serve as a think-and-do tank for governance innovation and offer policy, legal and technical guidance to public institutions deploying AI, among other objectives.[91]

Engagement to shape strategic and responsible AI

AI systems have the potential to radically reshape the interaction between citizens and their governments, and among citizens themselves. Key stakeholders, including members of the public, should have a say in how governments use and govern AI-based technologies. Involving citizens and stakeholders can lead to greater trust in and legitimacy of AI systems used by government, as well as to AI systems that better

reflect the needs of all (OECD, 2024[148]). Such efforts can help promote transparency, accountability and fairness in AI systems, preventing biases and potential harms.

Engaging citizens and stakeholders – such as scientists and engineers, affected communities, investors, companies or institutions — can enrich the understanding of issues related to AI technology, help policymakers anticipate problems of public acceptance and promote good communication (OECD, 2024[148]). This engagement is crucial for aligning the use and governance of AI with the goals and needs of society. It helps stakeholders to understand, question and influence how algorithms and AI governance mechanisms are designed and operated.

Early engagement allows for a comprehensive assessment of potential consequences and risks for various groups, fostering more inclusive and ethical development of AI systems and governance. This collaborative approach helps to identify and address concerns from diverse perspectives, ensuring that AI systems are designed and implemented in a way that is both responsible and beneficial to all stakeholders involved (OECD, 2021[112]). Citizens and stakeholders can be involved in different aspects of policymaking:

- **Agenda-setting**. In the United Kingdom, the Centre for Data Ethics and Innovation has been engaging citizens through the *Public Attitudes to Data and AI tracker survey*, now at its fourth wave, which can inform the government's approach to future policy development.[92]
- **Technology design**. In 2020, the French government launched PIAF, a collaborative initiative with citizens, academia and civil society to build databases in French language to train AI systems.[93] Greece's Pharos AI Factory is being shaped to be a central hub a central hub collaboration, knowledge sharing, resource pooling and joint project development among the public sector, academic institutions and private industry.
- **Technology assessment**. In the United States, the Expert & Citizen Assessment of Science & Technology (ECAST) Participatory Technology Assessment is bringing public perspectives to bear on critical government science and technology decisions. [94]
- **Regulation.** In 2024, the European Union opened a multi-stakeholder consultation on trustworthy general-purpose AI models under the AI Act,[95] as well as a call for expressions of interest to participate in the drawing-up of the first general-purpose AI Code of Practice.[96]

Governments have several options to consider in enhancing public engagement when shaping the strategic and operational development of AI. The sections below explore deliberative processes, such as **citizen assemblies, engagement with civil servants and the involvement of users in AI development.** In addition to civic engagement for government AI use and governance, governments can also use AI for civic engagement, which is discussed in depth in Chapter 5 (section on "AI in civic participation and open government").

Citizen assemblies

Citizens assemblies, also called citizens juries or panels, generally refer to a randomly selected group of people who are broadly representative of a community spending significant time learning and collaborating through facilitated deliberation to form collective recommendations for policymakers (OECD, 2020[149]).

A representative deliberative process is most suited to address issues such as values-based dilemmas; complex problems that require trade-offs and affect a range of groups in different ways; or long-term questions that go beyond electoral cycles (OECD, 2022[150]). AI's development and governance are well suited for such a process. AI involves ethical and societal discussions to decide on its uses in specific contexts (e.g. facial recognition) and can be considered as a technical issue with trade-offs between innovation and regulation. Most importantly, the adoption of AI technologies will certainly shape social interactions in the long term, with impacts that can span through generations.

As an example of a citizen assembly on AI, in 2024, the Belgian Presidency of the Council of the European Union convened a representative group of 60 Belgians to collect citizens' views on AI with the bloc (BeEU, 2024[151]). In another example, in the United States in 2023, a professor from Syracuse University partnered with the Center for New Democratic Processes (CNDP) to conduct the first national deliberative event on AI in the United States (Atwood and Bozentko, 2023[152]). International and sub-national examples of assemblies to shape AI governance also exist. In 2025 and 2026, the Global Coalition for Inclusive AI, a partnership between the Stanford Deliberative Democracy Lab and the Missions Publiques citizen participation consulting firm, will conduct deliberative assemblies that intend to reach more than 10 000 citizens in more than 100 countries. Follow-up events with decision-makers are planned to ensure the impact of the deliberative processes (Vergne and Siu, 2025[153]). Citizen participation is also a strong component of another emerging trend in AI governance: AI Localism, in which communities, including local governments, act to discuss and regulate the use of AI technologies according to their needs (Marcucci, Kalkar and Verhulst, 2022[154]). For example, in 2018, Mexico City's innovation lab, Laboratorio para la Ciudad de Mexico (LabCDMX), conducted a deliberative exercise build an Anticipatory Governance Framework for Mexico City on AI (Ramos, 2018[155]).

Engaging with civil servants and social partners

Engaging civil servants, who are at the frontline of public service delivery, is critical to the use of AI in government. Their roles and responsibilities are directly impacted by the introduction of AI technologies, and their insights and experiences are invaluable in shaping a responsible and effective AI use. As discussed in Chapter 1, the automation, creation or transformation of tasks due to the introduction of AI can bring about opportunities for improved efficiency and effectiveness. But those changes can also raise concerns about job security, work conditions and workers' right for these individuals.

The design and implementation of AI initiatives should be carried out in a manner that respects labour rights, promotes civil servants' well-being and uses their insights (OECD, 2023[118]). Transparent and inclusive dialogue with public servants and social partners, such as trade unions and employee associations, is important for achieving this. Workers should be informed about the objectives of AI initiatives, the potential impacts on their roles and the measures in place to mitigate any negative impacts. They should also be given opportunities to voice their concerns, provide feedback and contribute their insights to AI approaches.

Social dialogue and collective bargaining are essential for successful AI adoption in government (OECD, 2023[118]). They are key to building trust and effective collaboration among public servants and ensuring access to training to develop the skills and capabilities needed to work with AI. Social partners should also be involved, as they play a critical role in negotiating conditions for work.

Involving users in AI development

Involving end-users in AI development in government helps ensure that AI solutions are user-centric and effectively tackle real-world issues. Figure 4.5 shows key steps to understanding users and their needs. Governments can use research methods — such as reviewing existing evidence, conducting interviews and observing users — to develop a deep understanding of these aspects, thereby enhancing the relevance and acceptance of AI applications (OECD/UNESCO, 2024[29]). In line with OECD Good Practice Principles for Service Design and Delivery in the Digital Age, users can help identify insights for iterating the design of services, simplifying underlying procedures and increasing access for all user groups (2022[156]). Moreover, making the design and delivery of AI-driven services a participatory and inclusive process empowers users, giving them an active role in co-creating and co-designing public services. This can include implementing mechanisms to involve users in testing, iterating and improving the service, as well as conducting rigorous and ethically-designed experiments with user groups to help ensure the use

of AI has its intended effects and that any risks are identified on a small scale before scaling more broadly (see further discussion on "Creating spaces to experiment" above).

Figure 4.5. Key steps to understand the users and their needs in government AI developments

Source: (OECD/UNESCO, 2024[29]).

Collaborating across borders

Like other digital technologies, AI knows no borders. Its risks and impacts, as well as its potential positive uses, can be transnational.[97] Cross-border engagement and collaboration can be instrumental in bridging knowledge and development gaps among countries, tackling common and complex challenges, managing risks and implementing innovative policies and services. International co-operation can help in building government AI capacity across the globe and in specific regions, such as been seen in Latin America and the Caribbean (LAC) (OECD/CAF, 2022[30]). This can include sharing open algorithms, infrastructure and intergovernmental datasets, as well as conducting joint efforts for the responsible development of emerging technologies. OECD (2021[157]) work has identified three mechanisms governments are using to connect and collaborate in order to tackle issues that cut across borders between administrative entities or area, including in areas related to digital innovation and AI in government.

Cross-border governance bodies can address complex issues or those spanning the remit of multiple jurisdictions, such as seeking to integrate and harmonise AI approaches and interoperability. Governance bodies allow governments to coordinate and harness the collective efforts of actors divided by borders. This can be seen in the European Union's creation of the AI Office responsible for implementing, supervising and enforcing the AI Act. Beyond the EU sphere, the OECD Working Party of Senior Digital Government Officials (E-Leaders) and its thematic group on AI have been platforms for countries to work together to develop guidance and analytical products on AI in government, with the potential to propose relevant OECD non-binding legal instruments. Further, countries have been collaborating on international standards enforced by international bodies, such as ISO/IEC 42001 on AI management systems, relevant for public sector agencies as well as companies or non-profits. However, there is currently no evidence of formal cross-border governance bodies set up to specifically address government development and use of AI.

Second, countries are also using **innovative networks** tackling cross-border collaboration. Networks are horizontal, often informal and ground-up structures that allow for the organic convergence of ideas and expertise across borders. For instance, the European Public Administration Network (EUPAN) promotes knowledge sharing on AI.[98]

And third, some countries are exploring **emerging governance systems dynamics**, which are entirely new ways of working together across borders. For instance, governments have worked together to develop collective digital infrastructure and data sharing approaches to promote seamless operations internationally. The European Union is perhaps the most advanced in this area, as its Interoperable Europe Act, which came into force in April 2024, establishes a framework to enhance interoperability within public sector organisations, ensuring seamless cross-border services. Key elements include creating an

interoperability governance structure, promoting innovation and knowledge exchange, implementing regulatory sandboxes for testing solutions, and mandating interoperability assessments for public administration.[99]

A framework for trustworthy AI in government

When taken together, the policy measures discussed in this chapter form a *Framework for Trustworthy AI in Government* that can help governments align their actions in developing and adopting AI with the value-based principles and recommendations laid out in the OECD AI Principles (2024[158]). The framework outlines how governments can seize AI's promise of productivity, responsiveness and accountability by putting in place the right mix of enablers, safeguards and engagement mechanisms.

Figure 4.6 presents a visual representation of the framework and Table 4.1 details the policy questions and measures that underpin its elements.

Figure 4.6. OECD Framework for Trustworthy Artificial Intelligence in Government
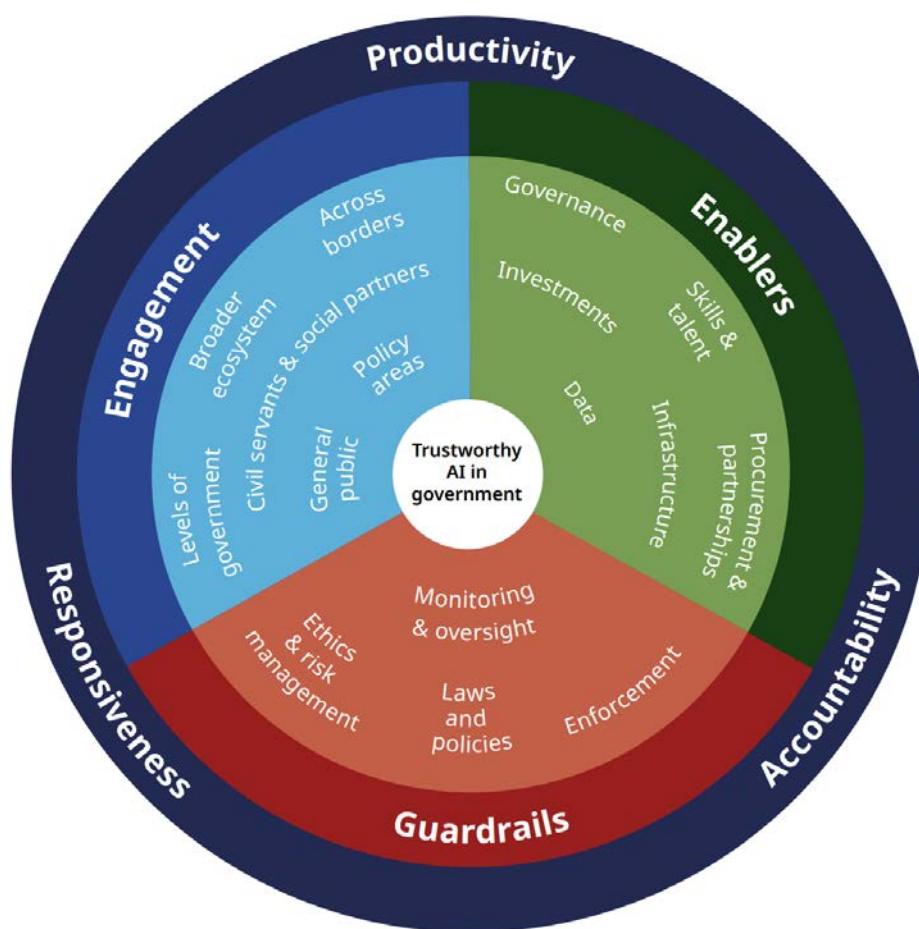
Source: (OECD, 2024[3]), as enhanced and finalised through the work conducted for this report.

Table 4.1. Policy questions and measures underpinning the Framework for Trustworthy AI in Government

Policy question	Policy measure	Description
What concrete policy actions and tools can governments develop to address existing challenges for a trustworthy use of AI in government?	Enablers	Policy actions and tools for areas where policymakers currently identify constraints and shortcomings, in order to establish a solid enabling environment and unlock the full-scale adoption of AI in government. These include governance, infrastructure, data, skills and talent, investments, procurement and partnerships.
	Guardrails	Policy tools that governments can consider developing for a responsible, trustworthy and human-centred use of AI in government. These may include non-binding instruments; laws and regulations; transparency and risk management instruments; or oversight (beyond the executive) and monitoring (within the executive) bodies.
Who should governments engage when developing and implementing the enablers and guardrails, as well as individual use cases, for the trustworthy use of AI in government?	Engagement	Different stakeholders to engage in building the foundations for a responsible use of AI in government. Various actors across government (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 opportunities and challenges related to the use of AI in government.
What impact does government strive to achieve when using trustworthy AI?	Impact	AI in government can help to increase productivity, responsiveness and accountability.

Source: (OECD, 2024[3]).

Future OECD work on these issues

The enablers, guardrails and engagement processes that comprise the OECD Framework for Trustworthy AI in Government serve as a strong foundation upon which governments can take a strategic and responsible approach to AI. However, one report cannot fully convey the complexities of this vast spectrum of activities needed to adopt rapidly evolving AI technology while managing both critical risks and significant implementation challenges. Future OECD work will address elements of the framework more in-depth, with actionable insights about how governments can put in place these foundations. For instance, a report on AI experimentation in government is already underway and will be released in the coming months.

Critically, governments need to be able to identify where to prioritise AI investments and resources based on various trade-offs when considering the potential benefits and risks of particular AI applications. The OECD has recommended (2024[3]) and continues to encourage governments to prioritise high-benefit, low-risk applications of AI, especially when building an initial level of maturity. However, most do not have the processes in place for holistic measurement of potential or realised results — efficiency of spend, quality of services, potential harms — that would allow them to make these determinations. This should be a priority for governments as a cross-cutting step that helps unlock the potential of AI, and it will be a focus of future OECD work.

References

Ada Lovelace Institute (2025), *Learn fast and build things: Lessons from six years of studying AI in the public sector*, Ada Lovelace Institute, https://www.adalovelaceinstitute.org/policy-briefing/public-sector-ai/. [23]

Ada Lovelace Institute (2021), *Algorithmic accountability for the public sector*, Ada Lovelace Institute (Ada), AI Now Institute (AI Now), and Open Government Partnerships (OGP), https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector/. [129]

African Union (2024), *Continental Artificial Intelligence Strategy*, [65]
https://au.int/en/documents/20240809/continental-artificial-intelligence-strategy.

Agency for Digital Government (2024), *A Common Danish Language Resource*, [46]
https://en.digst.dk/digital-governance/new-technologies/a-common-danish-language-resource/ (accessed on November 2024).

AI Hub (2024), "AI Data Finder", *aihub.or.kr*, [53]
https://www.aihub.or.kr/aihubdata/data/list.do?currMenu=115&topMenu=100 (accessed on November 2024).

AI Sweden (2024), *A shared digital assistant for the public sector*, [51]
https://www.ai.se/en/project/shared-digital-assistant-public-sector (accessed on October 2024).

Atwood, S. and K. Bozentko (2023), *U.S. Public Assembly on High Risk Artificial Intelligence 2023 Event Report*, https://www.cndp.us/wp-content/uploads/2023/12/2023-U.S.-PUBLIC-ASSEMBLY-ON-HIGH-RISK-AI-EVENT-REPORT-final.pdf. [152]

Australia DTA (2024), *APS trials generative AI to explore safe and responsible use cases for government*. [101]

Australian Government (2024), *Evaluation of the whole-of-government trial of Microsoft 365 Copilot*, https://www.digital.gov.au/initiatives/copilot-trial. [36]

BeEU (2024), *A citizen's view of artificial intelligence within the EU*, https://belgian-presidency.consilium.europa.eu/media/lzxauu4i/rapport-ia-en-web-v2.pdf. [151]

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. [4]

Brauneis, R. and E. Goodman (2017), "Algorithmic Transparency for the Smart City", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3012499. [133]

Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, European Commission, https://op.europa.eu/s/z4XY. [68]

Chilean Transparency Council (CPLT) (2024), *CPLT lanza recomendaciones de transparencia algorítmica en servicios públicos*, https://www.consejotransparencia.cl/cplt-lanza-recomendaciones-de-transparencia-algoritmica-en-servicios-publicos/. [126]

CNIL (2023), *« Bac à sable » données personnelles : la CNIL lance un appel à projets sur l'intelligence artificielle dans les services publics*, https://www.cnil.fr/fr/bac-sable-donnees-personnelles-la-cnil-lance-un-appel-projets-sur-lintelligence-artificielle-dans. [41]

CNIL (2023), *« Bac à sable » intelligence artificielle et services publics : la CNIL accompagne 8 projets innovants*, https://www.cnil.fr/fr/bac-sable-intelligence-artificielle-et-services-publics-la-cnil-accompagne-8-projets-innovants. [42]

CONPES (2025), *CONPES 4144, Política Nacional de Inteligencia Artificial*, https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4144.pdf. [31]

Corrêa, N. et al. (2023), "Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance", *Patterns*, Vol. 4/10, p. 100857, https://doi.org/10.1016/j.patter.2023.100857. [114]

Council of Europe (2024), *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence. [119]

Cover-Kus, H. (2024), *UK Government doubles down efforts to deploy AI across the public sector*, https://www.techuk.org/resource/uk-government-doubles-down-efforts-to-deploy-ai-across-the-public-sector.html. [97]

Deleware General Assembly (2024), *House Bill 333: An Act to Amend Title 29 of the Deleware Code Relating to the Artificial Intelligence Commission*, https://legis.delaware.gov/BillDetail/140866. [124]

Digdir (2024), *The Norwegian Resource Centre for Sharing and Use of Data*, https://www.digdir.no/datadeling/norwegian-resource-centre-sharing-and-use-data/2766 (accessed on November 2024). [60]

Dombo, F. (2023), *Public Sector Considerations for Successful Cloud Adoption*, https://news.sap.com/africa/2023/05/public-sector-considerations-for-successful-cloud-adoption/. [73]

EC (2025), *A pioneering AI project awarded for opening Large Language Models to European languages*, https://digital-strategy.ec.europa.eu/en/news/pioneering-ai-project-awarded-opening-large-language-models-european-languages (accessed on 10 March 2025). [87]

EC (2025), *Common European Data Spaces*, European Commission, https://digital-strategy.ec.europa.eu/en/policies/data-spaces. [55]

EC (2025), *The AI Continent Action Plan*, European Commission, https://digital-strategy.ec.europa.eu/en/library/ai-continent-action-plan. [67]

EC (2025), *Updated EU AI model contractual clauses*, European Commission, https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses. [103]

EC (2024), *Adopt AI study*, European Commission, https://op.europa.eu/s/z2tg. [96]

EC (2024), *GovTech: influencing factors, common requirements and recommendations*, European Commission, https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/document/govtech-influencing-factors-common-requirements-and-recommendations. [110]

EC (2024), *What factors influence perceived artificial intelligence adoption by public managers*, European Commission Joint Research Centre, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC138684/JRC138684_01.pdf. [8]

EC (2023), *EU model contractual AI clauses to pilot in procurements of AI*, https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai. [102]

Fink, K. (2017), "Opening the government's black boxes: freedom of information and algorithmic accountability", *Information, Communication & Society*, Vol. 21/10, pp. 1453-1471, https://doi.org/10.1080/1369118x.2017.1330418. [132]

France Élysée (2025), *Make France and AI Powerhouse*, https://www.elysee.fr/admin/upload/default/0001/17/d9c1462e7337d353f918aac7d654b896b77c5349.pdf. [64]

Frazier, K. (2025), *The Dangers of AI Sovereignty*, https://www.lawfaremedia.org/article/the-dangers-of-ai-sovereignty. [70]

G7 (2023), *Hiroshima Process International Code of Conduct for Advanced AI Systems*, https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems. [137]

G7 (2023), *Hiroshima Process International Guiding Principles for Advanced AI system*, https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system. [136]

Gartner (2024), *Compare AI Software Spending in the Government Industry, 2023-2027*, https://www.gartner.com/en/documents/5318363. [92]

Gob.cl (2025), *Latam GPT: Learn about the Latin American AI model developed in Chile*, https://www.gob.cl/en/news/latam-gpt-learn-about-the-latin-american-ai-model-developed-in-chile/ (accessed on 10 March 2025). [86]

Goodman, E. and J. Trehu (2022), *AI Audit-Washing and Accountability*, https://www.gmfus.org/news/ai-audit-washing-and-accountability. [143]

Government of Canada (2025), *AI Strategy for the Federal Public Service*, https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/gc-ai-strategy-overview.html. [13]

Government of Greece (2025), *The implementation of the new National Supercomputer "DAIDALOS" and the Data Center at the Lavrio TPP begins - The country acquires one of the most powerful computing systems in Europe*, https://mindigital.gr/archives/7331. [82]

Government of Greece (2024), *A Blueprint for Greece's AI Transformation*, https://foresight.gov.gr/wp-content/uploads/2024/11/Blueprint_GREECES_AI_TRANSFORMATION.pdf. [47]

Government of Ireland (2025), *Guidelines for the Responsible Use of AI in the Public Service*, https://www.gov.ie/en/department-of-public-expenditure-ndp-delivery-and-reform/publications/guidelines-for-the-responsible-use-of-ai-in-the-public-service. [19]

Government of Korea (2024), *A New Chapter in the Age of AI: Basic Act on AI Passed at the National Assembly's Plenary Session*, https://www.msit.go.kr/eng/bbs/view.do?bbsSeqNo=42&mId=4&mPid=2&nttSeqNo=1071. [122]

Government of New Zealand (2025), *Public Service AI Framework*, https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/artificial-intelligence/public-service-artificial-intelligence-framework. [18]

Government of Spain (2024), *Spain sets up the International Artificial Intelligence Advisory Council*, https://www.lamoncloa.gob.es/lang/en/gobierno/news/Paginas/2024/20240621-ai-advisory-council-meeting.aspx. [146]

Government of Spain (2024), *The Government approves the Artificial Intelligence Strategy 2024*, https://portal.mineco.gob.es/en-us/comunicacion/Pages/20240514-Gobierno-aprueba-Estrategia-IA-2024.aspx. [109]

Government of Switzerland (2025), *Strategy Use of AI systems in the Federal Administration*, https://www.bk.admin.ch/bk/en/home/digitale-transformation-ikt-lenkung/ikt-vorgaben/strategien-teilstrategien/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html. [14]

Government of the Dominican Republic (2024), *Estrategie Nacional de Inteligencia Artificial*, https://innovacionrd.gob.do/enia/. [12]

Government of Uruguay (2021), *AI Strategy for the Digital Government*, https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/publicaciones/ia-strategy-english-version/ia-strategy-english-version/ai-strategy-for. [15]

GPAI (2024), *Algorithmic Transparency in the Public Sector: A state-of-the-art report of algorithmic transparency instruments*, Global Partnership on Artificial Intelligence, https://gpai.ai/projects/responsible-ai/algorithmic-transparency-in-the-public-sector/algorithmic-transparency-in-the-public-sector.pdf. [125]

Gupta, A. et al. (2021), *The State of AI Ethics Report (Volume 5)*, https://arxiv.org/abs/2108.03929. [140]

Hassani, A. et al. (2022), *Escaping the Big Data Paradigm with Compact Transformers*, https://arxiv.org/abs/2104.05704. [79]

IEA (2025), *Energy and AI*, https://iea.blob.core.windows.net/assets/34eac603-ecf1-464f-b813-2ecceb8f81c2/EnergyandAI.pdf. [77]

IEA (2023), *Data Centres and Data Transmission Networks*, https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks. [76]

Jones, N. (2025), *Where AI Is Now: Smaller, Better, Cheaper Models*, https://www.scientificamerican.com/article/ai-report-highlights-smaller-better-cheaper-models. [80]

Komaitis, K., E. Ponce de León and K. Thibaut (2024), *The sovereignty trap*, https://www.atlanticcouncil.org/blogs/geotech-cues/the-sovereignty-trap. [69]

Lam, M. et al. (2023), "Sociotechnical Audits: Broadening the Algorithm Auditing Lens to Investigate Targeted Advertising", *Proceedings of the ACM on Human-Computer Interaction*, Vol. 7/CSCW2, pp. 1-37, https://doi.org/10.1145/3610209. [142]

Leslie, D. (2019), *Understanding artificial intelligence ethics and safety*, The Alan Turing Institute, https://doi.org/10.5281/zenodo.3240529. [116]

Letzing, J. (2024), *What is 'sovereign AI' and why is the concept so appealing (and fraught)?*, https://www.weforum.org/stories/2024/11/what-is-sovereign-ai-and-why-is-the-concept-so-appealing-and-fraught/. [63]

Long, D. and B. Magerko (2020), "What is AI Literacy? Competencies and Design Considerations", *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-16, https://doi.org/10.1145/3313831.3376727.

[90]

M Saiful Bari, Y. (2024), "ALLaM: Large Language Models for Arabic and English".

[48]

Marcucci, S., U. Kalkar and S. Verhulst (2022), *AI Localism in Practice: Examining How Cities Govern AI*, https://files.thegovlab.org/ailocalism-in-practice.pdf.

[154]

McKinsey (2024), *AI power: Expanding data center capacity to meet growing demand*, https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand.

[75]

Mehta, S., A. Rogers and T. Gilbert (2023), *Dynamic Documentation for AI Systems*, https://arxiv.org/abs/2303.10854.

[141]

METR (2023), *Responsible Scaling Policies (RSPs)*, https://metr.org/blog/2023-09-26-rsp/.

[144]

Metz, C. et al. (2025), *How A.I. is Changing the Way the World Builds Computers*, https://www.nytimes.com/interactive/2025/03/16/technology/ai-data-centers.html.

[78]

Monteiro, B., A. Hlacs and P. Boéchat (2024), "Public procurement for public sector innovation: Facilitating innovators' access to innovation procurement"*, OECD Working Papers on Public Governance*, No. 80, OECD Publishing, Paris, https://doi.org/10.1787/9aad76b7-en.

[100]

Montgomery, C., F. Rossi and J. New (2023), *A Policymaker's Guide to Foundation Models*, https://newsroom.ibm.com/Whitepaper-A-Policymakers-Guide-to-Foundation-Models.

[84]

Mulgan, G. (2019), *Intelligence as an outcome not an input*, http://www.nesta.org.uk/blog/intelligence-outcome-not-input.

[26]

Netherlands Court of Audit (2024), *Focus on AI in central government*, https://english.rekenkamer.nl/publications/reports/2024/10/16/focus-on-ai-in-central-government.

[127]

NIST (2023), *AI Risk Management Framework*, https://www.nist.gov/itl/ai-risk-management-framework.

[135]

Noor, E. and B. Kanitroj (2025), *Speaking in Code: Contextualizing Large Language Models in Southeast Asia*, https://carnegieendowment.org/research/2025/01/speaking-in-code-contextualizing-large-language-models-in-southeast-asia.

[88]

OECD (2025), "Effectively Managing Investments in Digital Government: An OECD Policy Framework"*, OECD Public Governance Policy Papers*, No. 76, OECD Publishing, Paris, https://doi.org/10.1787/5c324e91-en.

[93]

OECD (2025), *Enhancing Access to and Sharing of Data in the Age of Artificial Intelligence*, OECD Publishing Paris, https://www.oecd.org/en/publications/enhancing-access-to-and-sharing-of-data-in-the-age-of-artificial-intelligence_23a70dca-en.html.

[50]

OECD (2025), *OECD Regulatory Policy Outlook 2025*, OECD Publishing, Paris, https://doi.org/10.1787/56b60e39-en.

[120]

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. [74]

OECD (2024), "AI, data governance and privacy: Synergies and areas of international co-operation"*, OECD Artificial Intelligence Papers*, No. 22, OECD Publishing, Paris, https://doi.org/10.1787/2476b1a4-en. [40]

OECD (2024), "Assessing potential future artificial intelligence risks, benefits and policy imperatives"*, OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, https://doi.org/10.1787/3f4e3dfb-en. [138]

OECD (2024), "Digital public infrastructure for digital governments"*, OECD Public Governance Policy Papers*, No. 68, OECD Publishing, Paris, https://doi.org/10.1787/ff525dc8-en. [62]

OECD (2024), *Enabling Digital Innovation in Government: The OECD GovTech Policy Framework*, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/a51eb9b2-en. [108]

OECD (2024), "Fixing frictions: 'sludge audits' around the world: How governments are using behavioural science to reduce psychological burdens in public services"*, OECD Public Governance Policy Papers*, No. 48, OECD Publishing, Paris, https://doi.org/10.1787/5e9bb35c-en. [2]

OECD (2024), *Framework for Anticipatory Governance of Emerging Technologies*, OECD Publishing, https://doi.org/10.1787/0248ead5-en. [148]

OECD (2024), *G20 Compendium on Data Access and Sharing Across the Public Sector and with the Private Sector for Public Interest*, OECD Publishing, Paris, https://www.oecd.org/en/publications/g20-compendium-on-data-access-and-sharing-across-the-public-sector-and-with-the-private-sector-for-public-interest_df1031a4-en.html. [54]

OECD (2024), "Governing with Artificial Intelligence: Are governments ready?"*, OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, https://doi.org/10.1787/26324bc2-en. [3]

OECD (2024), *OECD Artificial Intelligence Review of Germany*, OECD Publishing, Paris, https://doi.org/10.1787/609808d6-en. [94]

OECD (2024), *OECD Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier*, OECD Publishing, Paris, https://doi.org/10.1787/a1689dc5-en. [83]

OECD (2024), *Public procurement*, https://www.oecd.org/en/topics/public-procurement.html. [106]

OECD (2024), *Recommendation of the Council on Artificial Intelligence*, OECD Publishing, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. [158]

OECD (2024), *The Digital Transformation of Norway's Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/1620e542-en. [99]

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. [52]

OECD (2023), "Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI"*, OECD Digital Economy Papers*, No. 349, OECD Publishing, Paris, https://doi.org/10.1787/2448f04b-en. [134]

OECD (2023), "AI language models: Technological, socio-economic and policy considerations"*, OECD Digital Economy Papers*, No. 352, OECD Publishing, Paris, https://doi.org/10.1787/13d38f92-en. [44]

OECD (2023), "Common guideposts to promote interoperability in AI risk management"*, OECD Artificial Intelligence Papers*, No. 5, OECD Publishing, Paris, https://doi.org/10.1787/ba602d18-en. [139]

OECD (2023), *Global Trends in Government Innovation 2023*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/0655b570-en. [130]

OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, https://doi.org/10.1787/08785bba-en. [118]

OECD (2023), "Regulatory sandboxes in artificial intelligence"*, OECD Digital Economy Papers*, No. 356, OECD Publishing, Paris, https://doi.org/10.1787/8f80a0e6-en. [35]

OECD (2022), *Going Digital to Advance Data Governance for Growth and Well-being*, OECD Publishing, Paris, https://doi.org/10.1787/e3d783b0-en. [56]

OECD (2022), "Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint"*, OECD Digital Economy Papers*, No. 341, OECD Publishing, Paris, https://doi.org/10.1787/7babf571-en. [71]

OECD (2022), "OECD Framework for the Classification of AI systems"*, OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, https://doi.org/10.1787/cb6d9eca-en. [58]

OECD (2022), "OECD Good Practice Principles for Public Service Design and Delivery in the Digital Age"*, OECD Public Governance Policy Papers*, No. 23, OECD Publishing, Paris, https://doi.org/10.1787/2ade500b-en. [156]

OECD (2022), *OECD Guidelines for Citizen Participation Processes*, OECD Publishing, https://doi.org/10.1787/f765caf6-en. [150]

OECD (2021), "Achieving cross-border government innovation: Governing cross-border challenges"*, OECD Public Governance Policy Papers*, No. 10, OECD Publishing, Paris, https://doi.org/10.1787/ddd07e3b-en. [157]

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, https://doi.org/10.1787/f161916d-en. [112]

OECD (2021), *Good Practice Principles for Data Ethics in the Public Sector - OECD*, OECD Publishing, Paris, https://www.oecd.org/gov/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm (accessed on 14 April 2025). [43]

OECD (2021), *OECD Report on Public Communication: The Global Context and the Way Forward*, OECD Publishing, https://doi.org/10.1787/22f8031c-en. [21]

OECD (2021), *Public Sector Innovation Facets: Mission-oriented innovation*, OECD Publishing, https://oecd-opsi.org/publications/facets-mission/. [9]

OECD (2021), *Recommendation of the Council for Agile Regulatory Governance to Harness Innovation*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0464. [121]

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. [7]

OECD (2020), *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave*, OECD Publishing, Paris, https://doi.org/10.1787/339306da-en. [149]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en. [34]

OECD (2019), "Data governance in the public sector", in *The Path to Becoming a Data-Driven Public Sector*, OECD Publishing, Paris, https://doi.org/10.1787/9cada708-en. [59]

OECD (2019), *Recommendation of the Council on Public Service Leadership and Capability*, OECD Publishing, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0445. [159]

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. [57]

OECD (2017), *Systems Approaches to Public Sector Challenges: Working with Change*, OECD Publishing, Paris, https://doi.org/10.1787/9789264279865-en. [1]

OECD (forthcoming), *Generative AI Experimentation in Government: A review of current practices*, OECD Publishing. [37]

OECD (forthcoming), *Mapping Relevant Data Collection Mechanisms for AI Training*. [38]

OECD.AI (2025), *The OECD-African Union AI Dialogue 2.0: From strategy to implementation*, https://oecd.ai/en/wonk/the-oecd-african-union-ai-dialogue-2-0-from-strategy-to-implementation. [115]

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. [30]

OECD/UNESCO (2024), *G7 Toolkit for Artificial Intelligence in the Public Sector*, OECD Publishing, Paris, https://doi.org/10.1787/421c1244-en. [29]

Pahlka, J. (2024), *AI meets the cascade of rigidity*, https://www.niskanencenter.org/ai-meets-the-cascade-of-rigidity/. [20]

Parankusham, K., R. Rizk and K. Santosh (2025), *LakotaBERT: A Transformer-based Model for Low Resource Lakota Language*, https://arxiv.org/abs/2503.18212. [49]

Pehlivan, C. and E. Valín (2023), *Spain establishes the EU's first AI supervisory agency*, https://techinsights.linklaters.com/post/102intj/spain-establishes-the-eus-first-ai-supervisory-agency. [145]

Peixoto, T., O. Canuto and L. Jordan (2024), *AI and the Future of Government: Unexpected Effects and Critical Challenges*, https://www.policycenter.ma/publications/ai-and-future-government-unexpected-effects-and-critical-challenges. [45]

Personal Information Protection Commission of Korea (2023), *Policy direction for safe use of personal information in the era of artificial intelligence [translated]*, https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=9083. [39]

Policy Lab Digital, Work & Society within the German Federal Ministry of Labour and Social Affairs (2024), *Guidelines for the Use of AI in the Administrative Work of Employment and Social Protection Services*, https://www.denkfabrik-bmas.de/fileadmin/Downloads/Publikationen/Guidelines_for_the_use_of_ai_in_the_administrative_work_of_employment_and_social_protection_services.pdf. [117]

Ramos, J. (2018), *Laboratorio Para La Ciudad (CDMX)*, https://actionforesight.net/laboratorio-para-la-ciudad-cdmx/. [155]

Ray, T. (2025), *Sovereign remedies: Between AI autonomy and control*, https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/sovereign-remedies-between-ai-autonomy-and-control/. [66]

Redapt (2023), *On-Premises vs. Cloud for AI Workloads*, https://www.redapt.com/blog/on-premises-vs-cloud-for-ai-workloads. [72]

Rudra, S. (2024), *OSI Calls Out Meta for its Misleading 'Open Source' AI Models*, https://news.itsfoss.com/osi-meta-ai/. [160]

Ryseff, J., B. De Bruhl and S. Newberry (2024), *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*, RAND, https://www.rand.org/pubs/research_reports/RRA2680-1.html. [24]

Ryseff, J. and A. Narayanan (2025), *Why AI Projects Fail*, https://www.rand.org/pubs/presentations/PTA2680-1.html. [25]

SDAIA (2025), *Government Cloud (Deem)*, https://sdaia.gov.sa/en/Services/Pages/Deem.aspx. [81]

Seger, E. et al. (2024), *Open-Sourcing Highly Capable Foundation Models: An Evaluation of Risks, Benefits, and Alternative Methods for Pursuing Open-Source Objectives*, https://www.governance.ai/research-paper/open-sourcing-highly-capable-foundation-models. [85]

The Alan Turing Institute (2023), *AI Skills for Business Competency Framework*, https://www.turing.ac.uk/skills/collaborate/ai-skills-business-framework. [91]

U.S. General Services Administration (2024), *AI Capability Maturity - Operational maturity areas*, https://coe.gsa.gov/coe/ai-guide-for-government/operational-maturity-areas/index.html#dataops. [61]

UCL IIPP (2019), *A Mission-Oriented UK Industrial Strategy*, https://www.ucl.ac.uk/bartlett/public-purpose/sites/public-purpose/files/190515_iipp_report_moiis_final_artwork_digital_export.pdf. [10]

UK Committee of Public Accounts (2025), *Use of AI in Government*, https://committees.parliament.uk/publications/47199/documents/244683/default/. [128]

UK DSIT (2024), *Global leaders agree to launch first international network of AI Safety Institutes to boost cooperation of AI*, https://www.gov.uk/government/news/global-leaders-agree-to-launch-first-international-network-of-ai-safety-institutes-to-boost-understanding-of-ai. [147]

UK DSIT (2020), *Guidelines for AI procurement*, https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement. [27]

UK Government (2025), *A blueprint for modern digital government*, https://www.gov.uk/government/publications/a-blueprint-for-modern-digital-government/a-blueprint-for-modern-digital-government-html. [6]

UK Government (2025), *Prime Minister: I will reshape the state to deliver security for working people*, https://www.gov.uk/government/news/prime-minister-i-will-reshape-the-state-to-deliver-security-for-working-people. [5]

UK Government Digital Service (2025), *Artificial Intelligence Playbook for the UK Government*, https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government/artificial-intelligence-playbook-for-the-uk-government-html. [22]

UK House of Commons (2024), *Governance of Artificial Intelligence (AI): Government Response*, https://committees.parliament.uk/publications/46145/documents/230927/default/. [98]

UK NAO (2024), *Use of articial intelligence in government*, https://www.nao.org.uk/wp-content/uploads/2024/03/use-of-artificial-intelligence-in-government.pdf. [16]

UNESCO (2024), *Consultation paper on AI regulation: emerging approaches across the world*, https://unesdoc.unesco.org/ark:/48223/pf0000390979. [111]

UNESCO (2023), *Ethical impact assessment. A tool of the Recommendation on the Ethics of Artificial Intelligence*, UNESCO, https://doi.org/10.54678/ytsa7796. [113]

University of Rome Sapienza (2024), *AI made in Italy: here is Minerva, the first family of large language models trained "from scratch" for Italian*, https://www.uniroma1.it/en/notizia/ai-made-italy-here-minerva-first-family-large-language-models-trained-scratch-italian (accessed on 10 March 2025). [89]

US IT Modernization Centers of Excellence (n.d.), *AI Guide for Government*, https://coe.gsa.gov/ai-guide-for-government. [28]

US OMB (2025), *Accelerating Federal Use of AI through Innovation, Governance, and Public Trust*, https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf. [17]

US OMB (2025), *M-25-22 Driving Efficient Acquisition of Artificial Intelligence in Government*, White House Office of Management and Budget, https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf. [105]

Valderrama, M., M. Hermosilla and R. Garrido (2023), *State of the Evidence: Algorithmic Transparency*, https://www.opengovpartnership.org/wp-content/uploads/2023/05/State-of-the-Evidence-Algorithmic-Transparency.pdf (accessed on August 2024). [131]

van Noordt, C., R. Medaglia and L. Tangi (2023), "Policy initiatives for Artificial Intelligence-enabled government: An analysis of national strategies in Europe", *Public Policy and Administration*, https://doi.org/10.1177/09520767231198411.   [95]

Vergne, A. and A. Siu (2025), *Global Coalition for an Inclusive AI*, https://global-ai-dialogue.org/.   [153]

Verhulst, S. and M. Sloane (2020), *Realizing the Potential of AI Localism*, https://www.project-syndicate.org/commentary/local-regulation-of-artificial-intelligence-uses-by-stefaan-g-verhulst-1-and-mona-sloane-2020-02?barrier=accesspaylog.   [33]

Vinnova (2022), *Designing missions*, https://www.vinnova.se/contentassets/1c94a5c2f72c41cb9e651827f29edc14/designing-missions.pdf?cb=20220311094952.   [11]

WAM (2024), *Hamdan bin Mohammed appoints 22 Chief AI Officers across government entities in Dubai*, https://www.wam.ae/en/article/b3kujwp-hamdan-bin-mohammed-appoints-chief-officers-across.   [32]

Werner, J. (2024), *New York Governor Signs AI Oversight Bill*, https://babl.ai/new-york-governor-signs-ai-oversight-bill.   [123]

World Bank (2025), *Global Trends in AI Governance: Evolving Country Approaches*, https://openknowledge.worldbank.org/entities/publication/a570d81a-0b48-4cac-a3d9-73dff48a8f1a.   [107]

World Economic Forum (2025), *AI Procurement Guideline*, https://www.weforum.org/publications/ai-procurement-in-a-box/ai-government-procurement-guidelines/ (accessed on 10 March 2025).   [104]

Notes

[1] See https://oecd-opsi.org/work-areas/systems-approaches.

[2] See https://oecd-opsi.org/work-areas/anticipatory-innovation and https://www.oecd.org/en/about/programmes/strategic-foresight.

[3] The context and use of "enablers" in this report are not the same as the "AI enablers" for AI systems generally discussed by AI policy and technical communities, which consist of data, algorithms and computational power ("compute").

[4] See https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023, https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024, and https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet, respectively.

[5] See (OECD, 2021[7]) for additional relevant material, including skills and competencies for digital government leadership. See also the OECD (2019[159]) *Recommendation on Public Service Leadership and Capability* for information on how countries can instil values-driven culture and leadership, and ensure skilled and effective public servants, and responsive and adaptive public employment systems.

[6] The OECD is engaged in supporting mission-oriented innovations through its Mission Action Lab, a joint initiative from the OECD Directorate for Science, Technology and Innovation (STI), Observatory of Public Sector Innovation (OPSI) in the Public Governance Directorate (GOV), and the Development Co-operation Directorate (DCD). See https://oecd-missions.org.

[7] Behavioural science is an interdisciplinary approach that encompasses the study of human behaviour and the design of strategies to change it. See https://www.oecd.org/en/topics/behavioural-science.

[8] See https://www.apsc.gov.au/initiatives-and-programs/workforce-information/research-analysis-and-publications/state-service/state-service-report-2023-24/fit-future/supporting-safe-and-responsible-use-artificial-intelligence.

[9] See https://www.digmin.dk/digitalisering/mere-om-digitalisering/digital-taskforce-for-kunstig-intelligens-.

[10] See https://www.mitre.org/news-insights/fact-sheet/federal-ai-sandbox and https://www.mitre.org/news-insights/news-release/mitre-establish-new-ai-experimentation-and-prototyping-capability-us.

[11] https://ai.gov.uk.

[12] https://alliance.numerique.gouv.fr.

[13] See https://govtech.justica.gov.pt/en/govtech-justica-english and https://www.ceei.es/legal&justiciatechlab/?r=vxpx6w6j70qr1jstvc6, respectively.

[14] Details provided in Spain's 2024 AI strategy, https://avance.digital.gob.es/es-es/notasprensa/paginas/20240514-gobierno-aprueba-estrategia-ia-2024.aspx.

[15] https://www.gov.uk/government/publications/the-magenta-book/guidance-on-the-impact-evaluation-of-ai-interventions-html.

[16] See https://www.gov.uk/guidance/repository-of-privacy-enhancing-technologies-pets-use-cases for a repository of use cases from different countries assembled by the UK.

[17] Other countries include Brazil, Canada, China, Egypt, Estonia, Finland, Germany, France, Hungary, Iceland, India, Israel, Japan, Korea, Latvia, Norway, Qatar, Singapore, Slovenia, South Africa, Spain, Thailand, Türkiye, the UK, and Viet Nam.

[18] See also https://www.ekt.gr/en/news/30774.

[19] See https://www.athenarc.gr/el/news/meltemi-proto-anoihto-megalo-glossiko-montelo-gia-ta-ellinika and https://www.athenarc.gr/en/news/llama-krikri-new-greek-ai-language-model-featured-kathimerini, respectively.

[20] See https://static.pib.gov.in/WriteReadData/specificdocs/documents/2022/aug/doc202282696201.pdf and https://www.indiatoday.in/technology/news/story/bhashini-ceo-amitabh-nag-talks-about-how-their-ai-

tool-is-bridging-indias-language-divide-2646101-2024-12-06; information supplemented by Government of India officials.

21 See https://interoperable-europe.ec.europa.eu/collection/open-source-observatory-osor/news/spanish-authorities-release-alia-ai-models.

22 https://www.canada.ca/en/treasury-board-secretariat/corporate/reports/2023-2026-data-strategy.html.

23 See https://www.gov.uk/guidance/national-data-strategy and https://www.dta.gov.au/digital-government-strategy, respectively.

24 See https://www.gov.br/governodigital/pt-br/capacitacao/capacita-gov-br.

25 See https://sbci.gov.ie/information-access/data-sharing-and-governance-act, https://eur-lex.europa.eu/eli/reg/2016/679/oj, https://eur-lex.europa.eu/eli/reg/2023/2854, https://eur-lex.europa.eu/eli/reg/2022/868/oj, https://eur-lex.europa.eu/eli/dir/2019/1024/oj, and https://ec.europa.eu/isa2/eif_en, respectively.

26 Digital Public Infrastructure (DPI) is defined as shared digital systems that are secure and interoperable and that can support the inclusive delivery of and access to public and private services across society. Such systems act as common digital building blocks that underpin government processes and services and enable digital government transformation at societal scale (OECD, 2024[62]).

27 https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados.

28 Model weights are "The variables or numerical values used to specify how the input (e.g. text describing an image) is transformed into the output (e.g. the image itself). These are iteratively updated during model training to improve the model's performance on the tasks for which it is trained" (Seger et al., 2024[85]). The use of "open-source" models for this report does not imply that such models are released under an open-source license approved by the Open Source Initiative (OSI), a nonprofit steward of the Open Source Definition (https://opensource.org/osd). OSI has criticised some companies that call their models open source because they only provide the weights for the model, and not other elements, such as the training data, code and training practices (Rudra, 2024[160]). Some argue that such models should be called "open weight" instead of "open source".

29 https://bigscience.huggingface.co/blog/bloom.

30 https://www.tech.gov.sg/products-and-services/for-government-agencies/productivity-and-marketing/vica.

31 See https://oecd.ai/catalogue.

32 https://indiaai.s3.ap-south-1.amazonaws.com/docs/empowering-public-sector-leadership.pdf. Information supplemented by Government of India officials.

33 See https://www.ipa.ie/ipa-overview/onelearning.2548.html and https://www.ypes.gr/ypourgeio-esoterikon-google-enarxi-epimorfotikis-drasis-gia-dimosious-ypallilous-me-thema-tin-techniti-noimosyni, respectively.

[34] See https://www.elementsofai.com.

[35] See https://www.canada.ca/en/government/system/digital-government/digital-talent-strategy.html.

[36] See https://chcoc.gov/content/skills-based-hiring-guidance-and-competency-model-artificial-intelligence-work.

[37] See https://innovadorespublicos.cl and https://innovadorespublicos.cl/events/652/.

[38] https://catalogue.csps-efpc.gc.ca/product?catalog=DDN3-E45&cm_locale=en.

[39] https://www.smart-city-dialog.de/en/about-us/international-smart-cities-network.

[40] https://alliance.numerique.gouv.fr.

[41] See: https://cnai.swiss/en/products/community-of-practice/

[42] See https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/datenpolitik/daten-und-ki/daten-und-ki-node.html. and https://www.digitale-verwaltung.de/SharedDocs/downloads/Webs/DV/DE/Transformation/akteurssteckbrief-beki.pdf.

[43] See https://www.digital.gov.au/policy/ai/pilot-ai-assurance-framework, https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles and https://www.digital.gov.au/policy/ai, respectively.

[44] See https://www.gsa.gov/about-us/newsroom/news-releases/technology-modernization-fund-seeking-proposals-fo-02082024.

[45] https://www.numerique.gouv.fr/services/guichet-financement-exploitation-valorisation-des-donnees/

[46] See https://www.numerique.gouv.fr/offre-accompagnement/pilotage-panorama/.

[47] https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai

[48] https://www.buyict.gov.au/sys_attachment.do?sys_id=e535e2ca935caa10438b39cdfaba103d.

[49] See also https://www.chilecompra.cl/2024/11/goblab-uai-presento-nueva-herramienta-para-una-ia-responsable-y-etica.

[50] https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467.

[51] See https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet and https://www.elysee.fr/en/sommet-pour-l-action-sur-l-ia/public-interest-ai.

[52] https://www.dataobservatory.net.

⁵³ See https://www.rtu.lv/en/university/for-mass-media/news/open/latvia-establishes-artificial-intelligence-centre and https://digital-skills-jobs.europa.eu/en/latest/news/latvia-establishes-artificial-intelligence-centre.

⁵⁴ See https://oecd.ai/ai-principles, https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, and https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system, respectively.

⁵⁵ https://docs.un.org/en/A/78/L.49.

⁵⁶ See https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023, https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024 and https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet, respectively.

⁵⁷ https://www.industry.gov.au/publications/voluntary-ai-safety-standard.

⁵⁸ See also https://oecd.ai/en/wonk/how-the-oecd-ai-policy-observatory-has-shaped-colombia-and-latin-americas-approach-to-ai-policy.

⁵⁹ https://aicm.ai.gov.eg/en/Resources/EgyptianCharterForResponsibleAIEnglish-v1.0.pdf.

⁶⁰ https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html.

⁶¹ https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/kuenstliche_intelligenz/kuenstliche_intelligenz-node.html.

⁶² https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper.

⁶³ See https://hyscaler.com/insights/bahrain-pioneers-ai-regulation and https://www.ita.gov.om/itaportal/Data/SiteImgGallery/2024731125545486/National%20Artificial%20Intelligence%20Policy.pdf, respectively.

⁶⁴ https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf.

⁶⁵ https://research-data.urosario.edu.co/file.xhtml?persistentId=doi:10.34848/YN1CRT/8OHRT0&version=1.0.

⁶⁶ https://www.gov.uk/algorithmic-transparency-records.

⁶⁷ See https://github.com/ombegov/2024-Federal-AI-Use-Case-Inventory for a centralised consolidation of AI use case inventories from across US federal government agencies.

⁶⁸ See https://www.algoritmospublicos.cl/repositorio, https://odap.fr/inventaire, and https://algoritmes.overheid.nl, respectively.

69 See https://algoritmeregister.amsterdam.nl/en/ai-register and https://ai.hel.fi/en/ai-register, respectively.

70 https://stip.oecd.org/stip/interactive-dashboards/policy-initiatives/2023%2Fdata%2FpolicyInitiatives%2F2329.

71 https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub.

72 https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592.

73 For useful context, the results of the Global Right to Information (RTI) Rating assesses the formal legal framework for the right to information for each country against a variety of categories (e.g. scope of applicability, requesting procedure, appeals process). See https://www.rti-rating.org/country-data.

74 See also the work of the of the OECD.AI Expert Group on AI Risk & Accountability (https://oecd.ai/site/risk-accountability).

75 See https://www.coe.int/en/web/portal/-/huderia-new-tool-to-assess-the-impact-of-ai-systems-on-human-rights.

76 See, for example, coverage of the topic in (OECD/CAF, 2022[30]).

77 See https://www.nist.gov/news-events/news/2024/05/nist-launches-aria-new-program-advance-sociotechnical-testing-and.

78 https://www.forgov.qld.gov.au/information-and-communication-technology/qgea-directions-and-guidance/qgea-policies-standards-and-guidelines/foundational-artificial-intelligence-risk-assessment-guideline.

79 The OECD is further exploring the concept of AI risk thresholds, as demonstrated by a September 2024 public consultation on the topic (https://oecd.ai/wonk/seeking-your-views-public-consultation-on-risk-thresholds-for-advanced-ai-systems-deadline-10-september).

80 See https://www.gov.uk/government/publications/emerging-processes-for-frontier-ai-safety/emerging-processes-for-frontier-ai-safety#responsible-capability-scaling.

81 Supreme Audit Institutions (SAIs) are public bodies responsible for the audit of government revenue and expenditure. By scrutinizing public financial management and reporting they provide assurance that resources are used as prescribed. See https://sirc.idi.no/about/what-are-sais for additional details.

82 https://digital-strategy.ec.europa.eu/en/policies/ai-board.

83 https://www.ombudsman.europa.eu/en/doc/closing-note/en/196934.

84 https://www.autoriteitpersoonsgegevens.nl/en/themes/algorithms-ai.

85 See, for instance, https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Public_Accounts_and_Audit/Publicsectoruseof AI/Report and https://committees.parliament.uk/work/6986/governance-of-artificial-intelligence-ai, respectively.

86 https://www.gov.ie/en/publication/a3be3-membership-of-the-ai-advisory-council.

87 https://data.govt.nz/leadership/advisory-governance/data-ethics-advisory-group.

88 https://bioethics.gr.

89 https://www.wa.gov.au/organisation/department-of-the-premier-and-cabinet/office-of-digital-government/western-australian-artificial-intelligence-advisory-board.

90 See https://ised-isde.canada.ca/site/ised/en/canadian-artificial-intelligence-safety-institute, https://digital-strategy.ec.europa.eu/en/policies/ai-office, https://www.economie.gouv.fr/actualites/la-france-se-dote-dun-institut-national-pour-levaluation-et-la-securite-de-lintelligence, https://aisi.go.jp, https://www.aisi.re.kr, https://t.ly/vCtd1, https://www.gov.uk/government/publications/ai-safety-institute-overview, and https://www.nist.gov/aisi, respectively.

91 https://indiaai.gov.in/article/india-takes-the-lead-establishing-the-indiaai-safety-institute-for-responsible-ai-innovation. Information supplemented by Government of India officials.

92 https://www.gov.uk/government/publications/public-attitudes-to-data-and-ai-tracker-survey-wave-4/public-attitudes-to-data-and-ai-tracker-survey-wave-4-report.

93 https://www.etalab.gouv.fr/ia-decouvrez-et-participez-au-projet-piaf-pour-des-ia-francophones.

94 https://ecastnetwork.org and https://issues.org/thinking-like-citizen-participatory-technology-assessment-weller-govani-farooque.

95 https://digital-strategy.ec.europa.eu/en/consultations/ai-act-have-your-say-trustworthy-general-purpose-ai.

96 https://digital-strategy.ec.europa.eu/en/news/ai-act-participate-drawing-first-general-purpose-ai-code-practice.

97 The OECD has issued a series of reports on "Achieving Cross-Border Government Innovation", which touch on challenges and success cases in a variety of areas, including AI. See https://cross-border.oecd-opsi.org.

98 See, for example, https://www.eupan.eu/wp-content/uploads/2024/12/HU_eNews-on-Examples-of-recently-launched-digital-public-service-innovations.pdf.

99 https://www.consilium.europa.eu/en/press/press-releases/2024/03/04/interoperable-europe-act-council-adopts-new-law-for-more-efficient-digital-public-services-across-the-eu.

5 Deep dive: The current status and future potential of AI in government

This chapter provides in-depth analysis regarding the current state of play, the untapped potential, the management of risks and challenges and the way forward for core functions of government. This consists of government policy functions (regulatory design and delivery, tax administration and public financial management), key government processes (civil service reform, public procurement, fighting corruption and promoting public integrity, policy evaluation, and civic participation and open government) and government services and justice functions (public service design and delivery, law enforcement and disaster risk management, and justice administration and access to justice). In doing so, it provides many examples of AI use cases that public administrations around the world are employing in pursuit of productivity, responsiveness and accountability. The contents of this chapter informed the findings of Chapter 2 on trends on AI in government.

Government policy functions

Government policy functions— which include tax administration, public financial management and regulatory design and delivery—are essential to effective governance. They shape legal frameworks and generate public revenue, as well as help ensure transparent, accountable use of resources, forming the foundation for stable and equitable policy outcomes. AI is being used in these functions to strengthen the backbone of effective and efficient public administrations.

AI in tax administration

For many years, tax administrations have been using AI to support activities across their operating model and have been actively exploring its potential to further enhance their operations, improve taxpayer services, increase tax compliance and prevent tax fraud. They are well placed to do this; collecting and analysing data are at the heart of many tax administration processes, which has facilitated the early adoption of rules-based AI systems. These systems have allowed tax administrations to analyse and extract insights from large volumes of data, enabling quicker identification of non-compliance and more precise targeting of limited resources towards high-risk cases.

This existing experience underpins investigations into how more advanced techniques can be deployed in a tax administration context. While technological advancements in AI offer significant potential to enhance and transform services, they can also amplify risks related to sensitive data processing and introduce new concerns. Therefore, tax administrations need to place a strong emphasis on privacy, security and the trustworthy application of AI.

Current state of play

Since 2022, the OECD *Inventory of Tax Technology Initiatives* (ITTI) survey has provided insight into how AI is being applied across tax administrations around the world. The latest survey shows that the main areas of application in OECD countries are detection of tax evasion and fraud, decision-making assistance and improving tax services (Figure 5.1).

Figure 5.1. AI deployments across OECD Members who use AI in tax administration

Note: 29 of the 38 OECD members report using AI in tax administration in the 2024 Inventory of Tax Technology Initiatives.
Source: OECD Data Explorer - Inventory of Tax Technology Initiatives 2024 (https://oe.cd/dx/ITTI2024).

Improving compliance and detection of evasion and fraud

To improve compliance and detect evasion and fraud, AI is frequently used to uncover hidden patterns of behaviour or new connections between transactions, assets or taxpayers within the data sources that a tax administration might already hold. As technological capabilities and internal expertise has grown, AI-driven techniques are now being used to analyse unstructured data sets, such as text from handwritten documents and information from public posts on social media, to further add depth to the analysis and

uncover connections about activities that could indicate tax evasion or non-compliance. These techniques are also used to support the detection or prevention of fraudulent attacks on the tax system by criminals who can use AI to create false claims for refunds on a potentially very large scale. As one example, Greece's Independent Authority for Public Revenue (IAPR) is leveraging AI to combat tax evasion by detecting compliance issues, automating complex procedures and enabling auditors to respond in real time.[1]

More recent use cases include applying pattern-detecting capabilities of tax administrations' AI systems to data from imaging systems to allow tax administrations to analyse maps and satellite imagery. By comparing satellite images over time, AI is being used, for example, to identify property alterations or new buildings or taxable assets that may not have been declared to the tax authority (Box 5.1). Greece has also developed an AI model for geo-locating swimming pools that have not been declared for tax obligations.[2]

Box 5.1. Detecting undeclared property development in France

To improve the process of detecting undeclared constructions or developments, the French tax administration (DGFiP) uses artificial intelligence and data enhancement based on aerial photographs taken by the National Institute of Geographic and Forestry Information (*Institut national de l'information géographique et forestière*, or IGN) as part of the Innovative Land (*Foncier Innovant*) project.

The algorithms make it possible to extract the outlines of buildings and swimming pools from the IGN's public aerial images, which can be consulted on the website geoportail.gouv.fr.

A computer process then verifies, from the owners' declarations to the DGFiP, whether the elements detected on the images are correctly taxed (for property taxes, in particular). A DGFiP official then systematically verifies each anomaly detected before any action is taken to remind or ultimately tax the owner of the property.

Source: (OECD, 2022[1]).

Assisting administrative decision-making processes

To make administrative processes more efficient, tax administrations are deploying AI to automate their internal processes, such as helping categorise and distribute similar cases or tasks. Similarly, AI is commonly used to differentiate simple, standardised cases, which can be resolved through automation, from complex cases, such as tax disputes, where the expertise and judgement of tax officials are most needed. This application allows a more efficient use of human and technological resources (Box 5.2).

Box 5.2. Improving decision-making in Brazil

Brazil had around USD 140 billion in assessed taxes waiting for decisions in administrative court tax appeals. It takes about six years for the appeal ruling. Under the AI Litigation Project, Brazil employed supervised machine learning (ML) when distributing groups of similar files to the same officers, in order to increase their speed in administering the file and taking decisions. The first trials, conducted with a sample of 2 000 manually labelled files, showed that supervised algorithms can attain sensitivity and specificity of over 80%.

Brazil also employed clustering algorithms to complete files either in full or in part. Additionally, a web-based report assistant tool is being developed to support officers' analysis and help in their goal of reusing blocks of text. The tool's resources include the presentation of suggested groups of files and paragraphs and the highlighting of sentences that turned out to be important for the clustering process. Officers can label files and paragraphs, and the labels are used to improve future suggestions.

Source: (OECD, 2022[1]).

Enhancing risk assessment processes

A significant driver of tax administrations' embracing AI is its ability to analyse data to score and prioritise risks. With AI, tax administrations can process large volumes of taxpayer data — from historical filings to transaction records and digital payment information — to build sophisticated risk models and then assign risk scores to specific types of behaviour or transactions. These models help tax administrations quickly identify possible non-compliance and focus resources on them. As the predictive capabilities of these systems has increased, tax administrations can now to identify potential issues more quickly, prevent them from escalating, and better support taxpayers in getting their filings right. To ensure reliability of these risk models, it is essential that the training data is accurate. Box 5.3 provides an example in Austria. As another example, in the Greek fuel market, IAPR has developed an analysis system that uses data from various sources to compose over 100 different risk analysis criteria.[3]

Box 5.3. Improving compliance checks with predictive analytics in Austria

Since 2014, the Austrian tax administration has applied machine-learning algorithms through the Predictive Analytics Competence Centre (PACC), a specialised unit within the Federal Ministry of Finance. Tasked with modernising risk management, the PACC works to improve tax collection, auditing and fraud detection. Organised into four subject areas (Predictive Analytics, Advanced Analytics, Tax Analytics, and Customs Analytics), the PACC addresses a broad range of challenges across the tax system.

In 2023, the PACC's risk models analysed around 6.5 million cases across income, corporate and value added tax sectors as well as customs transactions. The analyses detected instances of false reporting in employee tax assessments and identified fraudulent activities, resulting in additional tax revenues of approximately EUR 185 million. Nearly 27.5 million cases were also examined for compliance, with 375 000 cases flagged for further review due to implausible risk profiles. Advanced techniques, including decision trees, regression models and text mining, support both retrospective and real-time audits while ongoing projects seek to expand analytical capabilities to include, for instance, generative models.

Source: https://www.bmf.gv.at/en/press/press-releases/2024--New/August-2024/BMF-generated-around-EUR-185-million-in-tax-income-from-AI-in-2023-.html.

Improving taxpayer services

Helping taxpayers meet their obligations is a long-standing activity of tax administrations that often uses labour intensive service delivery methods, such as call centres or visits at physical tax offices. To help taxpayers faster with simple enquiries or processes and to free up resources, tax administrations have for many years adopted AI-based virtual assistant technology. Basic assistants use rules-based techniques that can direct taxpayers to relevant information, make payments or check their status. According to recent

ITTI data, virtual assistants that follow a set of pre-programmed rules when interacting with taxpayers are a very common use case for AI.[4] More recently, tax administrations have begun using more advanced AI techniques, such as large language models (LLMs) to enhance and personalise interactions with taxpayers. The example in Box 5.4 illustrates the range of services that are possible, and the growing levels of sophistication that allow more personalised responses. In another example, Greece's IAPR is developing an LLM-driven digital assistant to provide advisory support to taxpayers by answering questions during the submission of income tax returns.[5]

Box 5.4. Enhanced taxpayer services in Singapore

The Inland Revenue Authority of Singapore (IRAS) has employed end-user automation (i.e. tools that simplify things for service users), data and AI tools to deliver seamless and personalised taxpayer services. In 2021, IRAS launched a chatbot to more efficiently handle common queries on individual income tax, corporate tax, goods and services tax, property tax, stamp duty, withholding tax and employer tax matters. The chatbot is powered by the Singapore Government Virtual Intelligent Chat Assistant (VICA) platform.

The IRAS VICA chatbot enhances user engagement through a humanised user experience, featuring interactive elements such as carousels and infographics. In 2023, the chatbot was upgraded with the integration of a Large Language Model (LLM) engine, significantly improving its ability to understand the intent of taxpayers' queries.

Beyond answering queries, the chatbot also enables access to authenticated services. Taxpayers can conveniently check their outstanding tax balance, view payment plans, check payment status, cancel or reinstate payment plans and make tax payments via self-service payment terminals and mobile channels. These bot transactions save taxpayers about 10 minutes per transaction compared to traditional digital service channels. In the financial year 2024, the IRAS VICA chatbot handled about 70 000 transactional queries, potentially saving an estimated 11 666 taxpayer hours.

Source: (OECD, 2022[1]), https://www.iras.gov.sg/digital-services/others/iras-bot, https://www.tech.gov.sg/products-and-services/for-government-agencies/productivity-and-marketing/vica, Government of Singapore officials.

Assisting taxpayers in filing

Additionally, AI has facilitated the extensive pre-population of tax returns; algorithms and other tools examine tax administrations' extensive data sets and then place the data in the correct part of the tax returns. This can improve the accuracy of tax returns and simplify the filing processes, reducing the time taxpayers spend on filing their returns. In instances where taxpayers complete their returns, tax administrations are deploying pattern detection techniques that can find anomalies and errors and can prompt taxpayers to double check their returns.

Box 5.5. Preventing errors in Australian tax returns

The Australian Taxation Office (ATO) is employing advanced technologies to enhance the accuracy and efficiency of tax return submissions. These include real-time analytics, pre-filled forms and anomaly detection systems to assist taxpayers in meeting their obligations and reducing errors.

As part of its efforts, the ATO provides a pre-filling service that automatically populates individual tax returns with data sourced from employers, banks, government agencies and other third parties. This

---

includes information such as salary, bank interest, dividends and private health insurance details. Taxpayers are required to review and confirm the pre-filled data before submission. This system not only reduces administrative burdens but also improves accuracy by minimising manual entry errors.

Additionally, the ATO uses real-time prompts during the submission process to address potential anomalies. For instance, if a taxpayer's reported figures are flagged by a model as abnormal, , the system generates a message encouraging them to double-check their inputs. A message might state, for example: "You have not reported any interest income. Please review and ensure any interest earned has been reported." In 2023-24, more than 636 000 prompts were issued to individuals, helping to protect approximately AUD 78.9 million in revenue.

Source: (OECD, 2022[1]), https://aws.amazon.com/institute/future_of_tax_technology_ai_ml_cloud_security, https://www.austlii.edu.au/au/journals/JCULawRw/2024/5.pdf, https://www.ato.gov.au/individuals-and-families/your-tax-return/how-to-lodge-your-tax-return/lodge-your-tax-return-online-with-mytax/pre-filling-your-online-tax-return.

Managing risks and challenges

Whilst all deployments of AI in public administrations are generally accompanied by challenges, tax administration makes a particularly demanding environment for AI integration because of three critical factors:

1. Ensuring data protection and privacy: Tax administrations have access to a vast amount of sensitive taxpayer data and have the obligation to protect the privacy and confidentiality of this data. Governance and legal frameworks therefore need to be adapted to account for the use of AI.

2. Ensuring taxpayer rights: Taxpayers' ability to challenge tax decisions made by tax authorities mean that decision-making relying on AI systems and associated administrative processes need to be explainable, transparent and accountable.

3. Maintaining trust: Tax administrations rely heavily on taxpayers' voluntary compliance, which requires sustained trust in the impartiality and fairness of administrative processes. Taxpayers need to be able to trust the government's use of AI systems, continue to believe tax systems are fair and remain willing to comply with their tax obligations.

Associated risks

- Inadequate or skewed data in AI systems
- Lack of transparency and explainability

If AI systems rely upon **inadequate or skewed data**, it could lead to inaccurate or adverse outcomes for some individuals or groups. With regard to tax administration, this could result in inaccurate risk assessments and thereby the improper targeting of some individuals or groups for control measures (e.g. audits). See Box 5.6 for an example.

When adopting AI in tax administration, **inadequate and poor data quality** poses a significant risk to effective outcomes, which can undermine public trust. For AI systems to return reliable and trustworthy outcomes, the input dataset needs to be accurate, complete, and well-structured. Discrepancies in data formats, incomplete or mistyped records, outdated taxpayer data or data lacking the necessary context can lead to inaccurate predictions, flawed risk models and biased outcomes. Some machine learning (ML) systems do not require data to be as structured, but quality and accuracy are still critical. Incorrect inputs can make systems less efficient, undermine taxpayers' trust in government and have adverse consequences for taxpayers.

> **Box 5.6. Challenges with algorithmic errors in tax administration**
>
> The "Toeslagenaffaire" was a child benefits scandal in the Netherlands. Between 2005 and 2019, the Dutch Tax and Customs Administration (Belastingdienst) wrongfully accused approximately 26 000 families of fraudulently claiming childcare benefits, due to the use of flawed data and a skewed algorithm. Many of these families were flagged due to administrative errors, such as missing signatures, and disproportionately targeted families with dual nationality or migrant backgrounds. These families were forced to repay tens of thousands of euros in benefits, plunging them into severe financial hardship, with some losing their homes, jobs and marriages. In extreme cases, children were removed from their families. The ensuing scandal resulted in the collapse of the government.
>
> Following the Toeslagenaffaire, the Dutch Ministry of the Interior and Kingdom Relations tasked researchers at Utrecht University to develop Human Rights and Algorithms Impact Assessment (IAMA). In 2022 the Dutch House of Representatives approved a motion to "make it mandatory to conduct this impact assessment before using algorithms when algorithms are used to make evaluations or decisions about people."
>
> Source: (OECD, 2023[2]).

Challenges regarding the **transparency, explainability and interpretability** of complex AI systems can pose risks to rule of law and a taxpayer's legal recourse. The ability to challenge decisions is a fundamental component of tax administration processes. AI's "black box" nature for many ML systems can create a lack of transparency, explainability and interpretability that can impact on taxpayers' rights to understand how tax decisions are made, which in turn can limit their rights to dispute a decision.

To mitigate these risks, tax authorities need to invest in governance frameworks. These frameworks should ensure that AI is deployed in a way that respects the balance between the administration's tax collection and taxpayers' rights. Therefore, AI systems and associated operational processes in tax administrations need to be sufficiently transparent and explainable and operate with proper oversight and accountability mechanisms. These can include, for example, data governance (see Chapter 4, section on "Creating a strong data foundation"), including in terms of data cleaning, validation processes and regular updates to maintain the integrity and trustworthiness of the data used to inform AI-driven decision-making. Only with high-quality, reliable data can AI truly enhance tax administration by improving accuracy, compliance and operational efficiency for taxpayers.

Implementation challenges

- Skills gaps
- High costs of AI adoption and scaling
- Inflexible or outdated legal and regulatory environments

Taking advantage of the full range of opportunities that AI can bring to tax administration highlights three key challenges.

First, tax administrations report **fierce competition for skilled employees** who can develop and implement these technologies. Accompanying this is the requirement to train large numbers of existing employees in the use of this new technology and managing the associated changes in process.

Second, there are often **significant investment costs** associated with the deployment of this technology. The process for securing the budgets can be challenging, as the return on the investment may by uncertain

or over the long term. As well as helping to provide new services to taxpayers, this investment is also necessary to help defend the tax system against fraudulent attacks.

Finally, AI has the potential to **disrupt existing legal frameworks** that support many tax administration processes. Adapting these frameworks needs to be considered in parallel to the guidelines for ensuring the effective design of and deployment of AI. It is also worth noting that whilst countries may have general guidelines for AI deployment, they are often not tailored to the unique needs and specific complexities of tax administration, which needs further consideration.

Untapped potential and way forward

Tax administrations across the globe have already begun leveraging the potential of AI, deploying it in wide range of activities to improve the efficiency of their operations. AI holds to the potential not only to optimise tax administrations' existing operational models but also to transform them. OECD's (2020[3]) Tax Administration 3.0 Framework sets out a vision for digital transformation where digital technology is used to create seamless, event-based tax ecosystems. These allow for tax collection and reporting to be built directly into the natural systems of individual and business taxpayers.

To fully realise the transformative potential of AI, tax administrations should adopt a systemic approach to AI application. AI is not a standalone tool; it should be used an integrated component embedded in a broader tax ecosystem, encompassing regulatory frameworks, organisational structures, existing technical infrastructures and human expertise. It is crucial that tax administrations take into account the need for adjusting the rule-making process. For example, applying Rules as Code (RaC) approaches, tax administrations can encode tax laws and regulations into machine-readable formats, enabling AI systems to interpret and apply rules with greater accuracy and consistency (Mohun and Roberts, 2020[4]).[6]

Given that taxpayer trust is an essential part of a well-functioning tax administration, tax administrations should also ensure that AI deployed is trustworthy, transparent and accountable manner. To support tax administrations in this effort, the OECD Forum for Tax Administration is currently piloting a framework to support tax administrations in their deployment of AI systems. Building on the OECD principles for trustworthy AI, the framework outlines key considerations for each stage of the AI lifecycle, describing learning and experiences from a range of administrations from their own implementations. Through this ongoing co-operation, tax administrations can adopt a structured, risk-based approach to enhancing the use of AI in tax administration, which can also adapt to the changing wider context. The Forum on Tax Administration plans to publish its findings from this work in 2026.

AI in public financial management

Over the past few decades, public financial management (PFM) has regularly integrated new technologies, notably by adopting increasingly sophisticated Financial Management Information Systems (FMIS). As technology has advanced and the quality of data has improved, PFM organisations have adopted new technological approaches, such as through data analytics, business intelligence tools and robotic process automation (RPA). These have been the building blocks for the application of AI, which has generally been used to enhance and deepen existing capabilities, with AI systems and approaches being integrated into existing products and processes. In PFM, AI so far has generally been a continuation of the field's technological evolution rather than a technological revolution.

Current state of play

The current and envisioned uses of AI in PFM are mostly related to improving existing processes rather than fully re-envisioning and redesigning them (e.g. by removing any human intervention) or creating entirely new processes. AI is currently used as an assistant that provides high speed and low transaction cost in automating small, often mundane tasks for civil servants, and as an advisor that analyses historical or real-time data to forecast future events or behaviours to support civil servants' own analyses.

In this context, AI approaches, especially ML, have current applications for PFM in macroeconomic and macrofiscal forecasting[7] and supporting spending decisions; budget planning and monitoring; financial management, reporting and oversight; and engagement with external stakeholders.

Improving forecasting

AI can help address challenges with traditional financial and economic forecasting methods by enhancing the accuracy and timeliness of predictions, with AI systems able to outperform traditional economic forecasting models (Jung, Patnam and Ter-Martirosyan, 2018[5]).

AI's predictive capacities are also seen as an opportunity to develop nowcasting – identifying changes in near real time and extrapolating potential near-term futures. Nowcasting takes into account the very recent past and present to forecast the very near future state of economic indicators that are typically only determined after a delay and are subject to revision, such as gross domestic product (GDP) or inflation. Central banks of different countries are exploring the adoptions of AI systems to provide more accurate and anticipated predictions than traditional time series models, even when using unstructured data, such as New Zealand (Richardson, van Florenstein Mulder and Vehbi, 2019[6]), France (Blanchet and Coueffe, 2020[7]), and Peru (Tenorio and Perez, 2023[8]).

> ### Box 5.7. Forecasting GDP with explainable AI in Sweden
>
> The Swedish National Financial Management Authority (ESV) has developed an innovative GDP forecasting application leveraging explainable ML to enhance both accuracy and transparency in economic predictions. The model outperforms Sweden's pre-pandemic official forecasts and addresses a key limitation of traditional AI-based forecasting — its opacity — by visualising the impact of variables over time. This tool enables policymakers, researchers, and the public to generate reliable projections

while maintaining interpretability, supporting data-driven decision making. The initiative showcases how AI can be integrated into macroeconomic forecasting while ensuring accountability and trust.

Source: https://oecd-opsi.org/innovations/forecasting-gdp.

Facilitating spending decisions

To facilitate decision-making in PFM, AI is backed by technological advancements that have become mainstream. These include big data for analysing vast quantities of information from multiple sources, data analytics tools to drill down into specific financial categories and beneficiaries and evaluate the effectiveness of expenditures based on trends and patterns, and data visualisations to allow effective communication of complex information.

AI can be used to build on these foundations by identifying trends and patterns and grouping data points based on similarity or shared characteristics. For spending decisions specifically, AI can analyse historical budget execution data to identify underspending or overspending patterns, predict future spending needs based on key parameters (e.g. demographic changes), and evaluate programme effectiveness by linking expenditure data with outcome metrics. This offers opportunities to accelerate and improve analyses in ways that can automate or augment the work of humans. ML techniques that leverage unstructured data open opportunities for combining datasets previously unused for such exercises (Box 5.8).

Box 5.8. AI for public financial management in Korea

In 2022, South Korea developed and implemented dBrain+, an advanced financial management information system that leverages AI to analyse real-time economic, fiscal and financial data, optimising risk assessment and decision-making in public finance. Its key modules, the Korea Fiscal Information System (KFIS) and Korea Risk Assessment and Horizon Scanning (KORAHS), use AI-driven analytics to detect financial risks and support data-informed policy decisions. By centralising all national financial operations — from budgeting and fund management to debt oversight and performance evaluation — dBrain+ enhances efficiency, transparency and predictive capabilities across central and local governments.

A key strength of dBrain+ is its integration with 63 systems from 46 institutions, including the National Tax Service, Public Procurement Service and the Bank of Korea, enabling seamless coordination on contracts, tax collection and fund transfers. AI-powered analysis of this data in real time improves budget execution, accelerates financial reporting and supports the identification of risks, enabling better decision-making on fiscal policies and public spending. By providing tailored tools for different users — including civil servants, researchers and external stakeholders — dBrain+ strengthens accountability and modernises South Korea's approach to AI-driven fiscal governance.

Source: (Korea Fiscal Information Service (KFIS), 2023[9]), https://www.adb.org/sites/default/files/publication/928976/governance-brief-052-digital-transformation-tax-administration-rok.pdf.

Supporting budget planning and monitoring

AI has the capacity to support budget planning and monitoring processes by providing outputs that support the formulation of accurate expenditure baselines and costing of new policies. For instance, Australia's Department of Veteran's Affairs developed predictive systems and tools to help simulate future financial impacts of the policy decisions. These include the annual fiscal expenditure for each beneficiary as well

as their average years on benefits, which are used for costings, budget estimates and policy evaluation (Australian Government, 2020[10]).

Another promising domain of application for AI is the identification, monitoring and mitigation of fiscal risks through the analysis of large data sets. Government fiscal risks can arise due to a variety of causes, including unsustainable spending or investment levels, which need to be identified early to take preventive action. AI can support the identification of such risks, as seen in Box 5.9. In another example, Indonesia uses a system called AI Financial Advisor (AIFA) to process unstructured financial and performance data to provide analytics on subnational governments' fiscal performance in real time (Wisesa, 2023[11]).

Box 5.9. France uses AI in budget monitoring

For several years, France's tax agency (DGFiP) has implemented an AI-enabled "warning system" that aims to identify municipalities with financial difficulties, provide them with financial advice and proactively support the implementation of corrective measures.

This warning system was based initially on an algorithm using historical tax and financial data to score municipalities. More recently, DGFiP developed a predictive AI system designed to identify municipalities' financial difficulties earlier. The system was trained on data spanning four years to predict outcomes for the fifth year. The predictive system also relies on unsupervised clustering techniques to categorise municipalities with similar financial characteristics without predefined outcome examples.

In 2022, an experiment with the system covered 2 500 municipalities, with around 40% of these identified as facing financial difficulties. Of these, around 17% had not been detected by the previous algorithm. Additionally, around 35% of municipalities were identified with temporary, non-structural difficulties, highlighting the system's capacity to differentiate between permanent and transient financial problems.

Source: (French Public Finances General Directorate (DGFiP), 2024[12]).

Automating management, reporting and oversight activities

PFM and reporting activities involve important but sometimes repetitive tasks that are particularly well-suited for automation. AI techniques such as natural language processing (NLP) can be used to analyse digital images to extract information from documents (e.g. vendor information), identify and classify documents (e.g. invoices), perform document comparison (e.g. compare invoice and vendor information), or identify trends and patterns (e.g. internal controls on payment requests).

In France, for example, the French tax agency (DGFiP) (2024[13]) has developed an AI-based tool as part of the regular internal control processes that "automatises the selection of payment requests to be controlled [and] optimises the workload and the quality of controls performed". The Finnish Government Shared Services Centre for Finance and HR (Palkeet) established a Centre of Excellence for Robotic Process Automation (RPA). It focuses on developing and deploying RPA solutions across various financial and HR activities — such as the management of supplier information, balancing of accounting data and processing of financial transactions — and integrating AI into automation processes where complex decision-making or data processing are necessary (Palkeet, 2024[14]).

> **Box 5.10. AI-driven fiscal transparency in Brazil**
>
> Brazil's National Treasury (STN) is using AI to enhance fiscal transparency by classifying subnational government expenditures according to the international COFOG standard. Previously a manual and resource-intensive task, the adoption of AI — using ML models with Convolutional and Recurrent Neural Networks — has reduced classification time from 1 000 hours of human work time to just 8 hours, while achieving over 97% accuracy. This breakthrough led to the publication of the Expenditure by Function of the General Government report in 2024, a milestone in Brazil's fiscal statistics.
>
> Building on this success, STN is now expanding AI applications to new areas, including the classification of climate-related expenditures. In collaboration with the Inter-American Development Bank (IDB), Brazil is strengthening its capacity to assess the fiscal implications of climate change. By pioneering AI-driven public finance management, Brazil sets an example for other nations seeking to modernise fiscal statistics and enhance transparency in an increasingly complex economic landscape.
>
> Source: https://blog-pfm.imf.org/en/pfmblog/2024/12/ai-is-enhancing-fiscal-transparency-in-brazil.

In addition to strengthening reporting, the development of targeted verifications to identify errors (e.g. improper payments) and fraud (e.g. identity theft) have become a major objective for many governments. This is especially true in the wake of the COVID-19 pandemic, which exposed vulnerabilities in several countries' payment systems. AI's capacity for identifying trends and patterns can help with this. For instance, during the pandemic, the Danish Business Authority developed AI-based controls for aid applications from companies for various support schemes (van Noordt and Tangi, 2023[15]).

Facilitating engagement with stakeholders and users

As discussed throughout this chapter and synthesised in Chapter 2, chatbots powered by NLP and language models are increasingly employed in government to directly provide services. As related to PFM, the United Arab Emirates (UAE) has developed U-Ask, a unified AI-powered chatbot for government services that can also be used to answer simple fiscal reporting questions[8]. In Mexico, the government has introduced an AI virtual assistance tool as part of its Intelligent Support Platform, designed to guide users through government programmes and supports (2023[16]). The tool provides information on benefits, eligibility and application processes for individuals, businesses and local governments, utilising a simple keyword search or personalised questionnaires to tailor the information to the user's profile.

Managing risks and challenges

Associated risks

- Lack of transparency and explainability
- Inadequate or skewed data in AI systems

Due to their black box nature, AI-based systems that have the best forecasting outputs represent both a step forward in accuracy and a step back in fiscal transparency (Jung, Patnam and Ter-Martirosyan, 2018[5]). This **lack of transparency** makes it difficult for governments to verify the decision-making processes of these models, which is crucial for accountability and regulatory compliance. Consequently, this limitation has prompted governments to prioritise the use of simpler AI systems to improve human modelling and sensitivity analysis.

While governments and PFM organisations have become increasingly adept at identifying AI-related risks and challenges, many are still developing comprehensive frameworks and practical approaches to manage these risks. Their efforts focus on establishing governance structures, building technical capacity and creating clear protocols for AI deployment in public financial systems.

Governments are also working to develop methods to "unbox" AI systems and make their reasoning more transparent, explainable and interpretable — all important conditions for using these systems in PFM. For instance, as discussed above (Box 5.7), the Swedish National Financial Management Authority (ESV) has developed an application for analysing the impact that each data variable has on the prediction of the "black-box models", as part a wider work programme to integrate AI in the financial management of Swedish government (Boström et al., 2020[17]).

AI use cases show that **ethical risks like incomplete or insufficient data and skewed algorithms** can be significant in the field of PFM. AI can amplify patterns of inequality embedded in financial data, leading to financial exclusion of perceived high-risk individuals (Crisanto et al., 2024[18]). This phenomenon is observed in the banking industry, where biased credit distribution may perpetuate discriminatory lending practices (Bailey, 2023[19]; Klein, 2020[20]). In the realm of public financial management, biased algorithms can similarly affect the distribution of public funds, social benefits and access to government programs, perpetuating existing inequalities and hindering fair treatment. While more of an automated decision-making system than true AI, Australia's Robodebt scheme illustrates problems that can arise as a result of algorithmic errors if not caught and addressed by humans (Box 5.11).

Box 5.11. The Robodebt scheme: Challenges with collecting improper payments

Australia's Robodebt scheme, introduced in 2016, was an automated debt recovery programme designed to identify and recoup welfare overpayments. It replaced a manual process with a data-matching algorithm that compared fortnightly income data reported to Centrelink, the agency responsible for social security payments, with averaged annual income figures from the Australian Taxation Office (ATO). Discrepancies were flagged as overpayments, and debt notices were automatically issued without human verification. This "income averaging" method ignored fluctuations in actual earnings, often generating false debts. The system also reversed the burden of proof, requiring recipients to provide historical pay records to contest debts — a demanding task for many. Over its operation, the scheme issued 470 000 incorrect debt notices totalling EUR 775 million, causing widespread distress and financial hardship.

The scheme's calculations were declared unlawful in 2019. A Royal Commission was set up in 2022 to enquire into the establishment, design and implementation of the Robodebt scheme; the use of third-party debt collectors under the Robodebt scheme; concerns raised following the implementation of the Robodebt scheme; and the intended or actual outcomes of the Robodebt scheme. The Royal Commission issued a report in 2023, which discussed impacts the scheme had on recipients, including those related to income withholdings and garnishees, emotional and psychological effects, and the loss of faith in the government. Robodebt exemplifies the risks of automating complex social systems without adequate human oversight or rigorous testing. The fallout included significant legal settlements and calls for stricter regulations on the use of algorithms in public policy. While the Royal Commission's findings were not necessarily representative of the Australian Government's views, the government agreed, or agreed in principle, to 56 of the commission's 57 recommendations.

The Robodebt scheme — which used automated data-matching, income averaging and overpayment calculation — can be described as an automated decision-making system. While the scheme did not leverage AI, it helps to illustrate issues in governance, human oversight and algorithmic design.

Source: (OECD, 2023[2]), https://robodebt.royalcommission.gov.au, https://ministers.dss.gov.au/media-releases/13091.

Implementation challenges the process in a way to minimise

- Matching problems to AI solutions
- Inflexible or outdated legal and regulatory environments
- Outdated legacy information technology systems
- Lack of high-quality data and the ability to share it
- Skills gaps
- Lack of actionable frameworks and guidance on AI usage

One challenge is to **match PFM needs and AI technologies.** Most finance ministries that have already implemented AI projects in PFM emphasise the importance of mapping processes and activities to pinpoint areas of inefficiency and potential efficiency gains as a prerequisite to deploying AI.[9] Once this initial phase is complete, the next step is to assess the suitability of AI technology, or other technologies, for integration to help address them.

Despite AI's capacity to summarise or draft text using language models, finance ministries have been cautious in rolling out new technologies in fiscal reporting (e.g. the automatic production of fiscal reports). This may be due to **concerns over accuracy and whether AI complies with current regulations.** There may also be concerns over where responsibility lies when AI is used, and what its use means for people in positions of responsibility, such as external auditors, and people who use the reports, such as lawmakers.

Information technology (IT) systems are crucial for finance ministries to be able to take advantage of AI opportunities. Yet many OECD countries say they are **locked into legacy technologies** that are significantly fragmented, often outdated and lack the necessary infrastructure and compatibility to integrate advanced AI functionalities. For example, centrally managed FMIS systems are more than 10 years old in most OECD countries (OECD, 2024[21]). These technologies are not specific to OECD countries and are holding back use of AI in PFM across the world (Rivero del Paso et al., 2023[22])

As with any IT system, the quality of output obtained from an AI system depends on the quality of inputs. Finance ministries also indicate that **fragmented data, coupled with restrictions on data sharing**, frequently impedes the initiation of AI projects (see Chapter 4, section on "Creating a strong data foundation").[9] These challenges underscore the need for improved data management practices and policies that facilitate more effective data accessibility and sharing. Accordingly, various OECD countries plan major upgrades to their FMIS, also recognising the need for stronger data foundations (Figure 5.2).

Figure 5.2. Objectives for FMIS upgrades in OECD countries, 2022
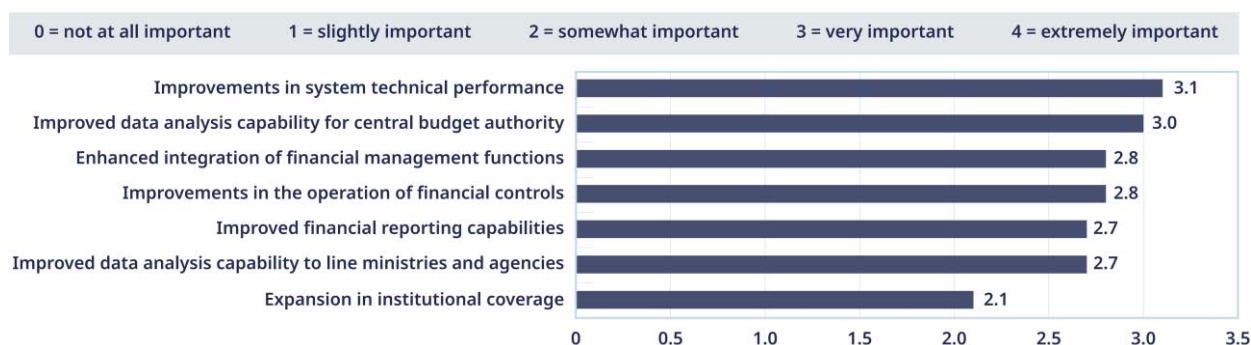
Note: Refers only to countries currently undertaking major development or replacements of their central FMIS (18 countries). Ratings present the average level of importance assigned to each objective on a scale of 0 to 4 by all respondents. Data for Chile, Colombia, Israel, Mexico, Slovenia and the United States are not available.
Source: (OECD, 2022[23]), Question 24.

Because PFM is a highly technical policy area, implementing AI systems requires substantial training that involves human intervention and supervision (and feedback) from PFM experts. A significant challenge is that while PFM organisations can effectively identify AI-related risks, many lack the **specialised staff skills and institutional capacities** needed to develop and implement necessary risk management frameworks. Further, it is tempting to overestimate AI's capabilities. AI systems produce outputs that are based on probabilities, which means they are by nature uncertain. They can produce data outputs that are wrong or texts that sound highly authoritative but are incorrect ("hallucinations"). Therefore, no matter how much training systems may receive, PFM specialists need to be able to exert critical judgment when using the outputs they generate. Such oversight requires specialists that combine technical PFM skills and basic understanding of how AI works.

Ensuring transparency and explainability in AI-driven decisions **requires robust frameworks and standards that govern and oversee AI processes**, including the traceability and accessibility of training datasets (e.g. data provenance, data records) and, when possible, the source code for AI algorithms. However, such frameworks and standards have until recently lacked in many OECD countries. Traditional oversight bodies, such as supreme audit institutions (SAIs) and independent fiscal institutions, are beginning to adapt their methodologies and develop new skills to effectively oversee AI-driven fiscal processes (Box 4.8). Additionally, identifying and implementing safeguards against potential misuse or over-reliance of AI in fiscal governance is crucial as AI systems become more prominent in decision-making processes. PFM specialists need to address these issues to shape a new fiscal governance framework that exploits AI's potential while maintaining transparency, accountability and integrity.

Untapped potential and way forward

Finance ministries are currently adopting a cautious approach to AI, prioritising task automation and predictive applications over more prescriptive AI. While predictive AI focuses on forecasting outcomes, prescriptive AI goes further by suggesting courses of action to achieve desired goals or mitigate risks. However, a systematic application of prescriptive AI could profoundly impact the roles and responsibilities within PFM systems, potentially altering the activities of fiscal stakeholders and oversight bodies.

Any shift from human judgment to system-based outcomes in fiscal decisions necessitates a risk-based re-evaluation of accountability, and the assignment of roles and responsibilities in an increasingly automated environment. This also requires considering how AI could alter the letter and spirit of PFM institutions and processes, and how it could reshape the functions of fiscal stakeholders, including external

oversight bodies such as SAIs and independent fiscal institutions. In this context, questions regarding the future of transparency and accountability need to be addressed.

- What mechanisms will be needed to ensure transparency in automated PFM decisions? As forecasting and budget planning and monitoring could increasingly be conducted by AI, there should be robust frameworks to ensure transparency and explainability regarding governments' use of AI systems and their underlying data. This involves creating and implementing standards that govern AI use and oversight.

- How will the roles of traditional oversight bodies evolve as AI is adopted? SAIs and independent fiscal institutions will need to adapt their methodologies to effectively oversee and audit AI-powered fiscal processes. This might include developing data science and AI training for staff members or updating audit processes to incorporate AI-specific considerations, such as data integrity and completeness.

- What new safeguards will be necessary to protect against misuse of AI in fiscal governance? As AI systems play a more prominent role in fiscal decision-making, identifying possible types of intended or unintended misuses and safeguarding against them becomes paramount.

By addressing these questions, PFM specialists can help shape a new fiscal governance framework that accommodates the full innovative potential of AI, while safeguarding the key principles of PFM.

Finance ministries' AI efforts need a coordinated approach and should integrate lessons from other governmental projects. This could greatly minimise risks of failure and implementation challenges, while enhancing the quality of outcomes. For instance, community of practice on AI projects can help to prevent common and avoidable mistakes and mitigate risks from inherently complex AI projects. These risks stem, for example, from the involvement of multiple stakeholders with diverse interests or lack of understanding of technologies and systems.

While many financial management agencies identify a potential for large-scale productivity gains from AI, public studies on feasibility and costs are rare and assessments on results and impacts remain anecdotal. Evidence on costs and impacts are either not collected or not publicly available due to the provisional nature of projects still in pilot or early development phases.

Ideally, the results and impact of AI use in PFM should be assessed by using evaluation frameworks to track costs of projects and key performance indicators from completed projects, including cost savings, effectiveness and efficiency gains, error reduction and compliance enhancement. This would involve, among other things, monitoring full costs of projects, comparing metrics from before and after AI implementation, conducting stakeholder surveys for satisfaction, and analysing data to see how AI outcomes match with fiscal policy goals.

Due to the lack of impact data, external oversight bodies recently called for greater transparency in AI projects across government; they said that greater scrutiny and evidence collection is required alongside substantial investment (see Chapter 4, sections on "Investing purposefully" and "Empowering oversight and advisory bodies to guide responsible AI").

Establishing robust project selection and evaluation frameworks within finance ministries is critical. These should be aligned with government-wide frameworks. They should track not only cost and performance indicators but also the qualitative impacts of AI, such as error reduction and compliance enhancement. This is needed, as multiple projects will likely compete in the future for limited investment resources.

To scale up their ambitions progressively and safely, finance ministries could adopt a sequenced approach to the introduction of AI technologies. This is well illustrated by the case of The Finnish Government Shared Services Centre for Finance and HR (Palkeet), as discussed above, which started with "small" uses of RPA and is now moving towards hyper-automation with ML. According to Palkeet, starting small with their

advanced digitalisation journey also helped increase the acceptability for civil servants of more sophisticated AI technologies (Palkeet, 2024[14]).

AI in regulatory design and delivery

AI has the potential to contribute to significant improvements in regulatory design and delivery. It can allow governments to tailor regulatory strategies for better economic, societal and environmental outcomes, while effectively addressing relevant challenges and concerns such as accountability, transparency and regulatory burdens. To achieve this, governments need to move away from the traditional "regulate-and-forget" approach of regulatory policymaking and adopt an "adapt-and-learn" approach (OECD, 2024[24]).[10] Digital technologies, including AI, play a pivotal role in this shift, not only by extending the regulatory toolbox but also by enabling more innovative, effective and efficient rule making through data-driven design, decision-making and enforcement (OECD, 2021[25]).

Current state of play

With regards to improving regulatory *design*, AI offers policymakers a number of significant opportunities. Regulatory systems can be complex, with thousands of pieces of legislation across numerous line ministries and agencies. AI can help navigate an existing stock of regulations and can analyse vast and complex datasets to identify gaps, overlaps and patterns in regulatory frameworks. This can enable more informed and targeted design decisions to improve outcomes. AI can also automate routine tasks, which can create efficiencies by streamlining process such as policy analysis, regulatory impact assessment (RIA), and legal drafting. It can strengthen stakeholder engagement by facilitating simplified regulatory understandings and analysing public consultations. AI can also enable anticipatory analysis and experimentation that can inform risk management, improving the design of future-ready regulatory frameworks.

Regarding regulatory *delivery*, AI can help improve the performance of delivery agencies. They can, for instance, use AI algorithms to optimise inspection resources, improve safety and reduce regulatory burdens for businesses. Regulators have been using AI to improve the activities and processes of these institutions to better protect public interests and resource efficiency (OECD, 2021[25]). Specific approaches include:

- Enhancing risk-based approaches with data analysis that help increase the accuracy and parameters of risk assessments to improve how inspections are targeted;
- Improving efforts to monitor and detect non-compliance, such as on social media platforms, which enhances oversight efficiency of regulatory agencies; and,
- Enhancing the diverse activities of economic regulators to better understand and oversee their sectors and markets.

Drafting regulations and related documents

AI systems that use NLP, such as LLMs, offer significant opportunities for simplifying and enhancing the written aspects of regulatory governance. Governments can use such systems to generate drafts from templates and existing regulation, ensuring that standards are met while saving time and resources. In the United States (US), for example, the state legislature of California became the first in the country to use AI to draft a resolution in 2023 (Tribune News Service, 2023[26]). This was closely followed by Costa Rica, which used ChatGPT drafts law to regulate AI (Guio and Müller-Daubermann, 2024[27]). AI can also be

used to cross-reference new drafts with current laws, identifying conflicts and reducing human errors, leading to simpler, more consistent and accessible regulations (Box 5.12).

Box 5.12. Using AI for legislative drafting and responses in the United Kingdom

The UK Government Incubator for Artificial Intelligence (i.AI) has developed two AI-powered systems, Lex and Parlex, to enhance legislative drafting and policy formulation.

Lex: This system aims to improve the legislative drafting process by providing advanced AI tools to navigate, explain and interrogate UK law. Key features include semantic search capabilities that allow users to locate contextually relevant legislative materials efficiently, and AI-assisted drafting tools that generate explanatory notes for government bills, reducing manual effort and increasing precision in legal language. The system was also designed with a deep understanding of UK-specific legal terms, allowing it to accurately capture the nuances of legal terminology, thus promoting innovation and collaboration in the legal sector.

Parlex: This AI system is designed to assist policymakers by forecasting parliamentary reactions to proposed policies. By analysing historical parliamentary records, Parlex provides insights into how MPs might respond to new policies, enabling officials to develop effective strategies for policy implementation. For example, it can perform a "parliamentary vibe check" on potential laws, predicting support or opposition among MPs based on past debates. This tool helps policymakers understand the political climate and anticipate challenges or support for policies before formal proposals are made.

Both Lex and Parlex exemplify the United Kingdom's commitment to integrating AI into public services, aiming to enhance efficiency, accuracy and strategic planning in legislative and policy development.

Source: https://ai.gov.uk/projects/parlex-and-lex, https://www.thetimes.com/uk/politics/article/parlex-ai-to-advise-ministers-on-how-policies-will-be-received-99txwlwph.

Enhancing the agility of regulatory assessments

Ex ante regulatory impact assessment (RIA) and ex-post evaluation are foundational elements of sound regulatory governance. They should not be seen as discrete requirements to be conducted successively, but rather as mutually complementary tools embedded in the policy cycle to inform the appropriate adaptation of regulatory (or alternative) approaches (OECD, 2024[24]; 2021[28]).

Regulators can use AI to enhance each of these elements, using AI to improve the speed and accuracy of assessment, while minimising the burden of such activities. This can allow for more frequent assessment of regulation to help create adaptive and future-fit regulations that remain relevant and effective in a rapidly changing environment. This use of AI may be especially necessary for ensuring appropriate governance of AI. AI experts note as a top risk that governance mechanisms and institutions are unable to keep up with the rapid pace of technological developments (OECD, 2024[29]).

As an initial step, AI can help estimate the burdens of potential new regulations and other policies, which can have a bearing on how easily companies can accommodate regulatory changes and how likely they are to comply. This naturally affects the success of any regulatory measure. It is also an important consideration for regulators comparing different governance approaches and their associated cost-benefit trade-offs for more efficient and economically sound policymaking (Box 5.13).

Box 5.13. Developing AI for measuring regulatory compliance costs in Germany

In Germany, the Service Centre for Better Regulation in the Federal Statistical Office is developing a ML system to help estimate compliance costs to support regulatory impact. This approach involves identifying passages of legal drafts that influence compliance costs, using AI for web scraping of legal texts. Then it uses AI to predict which new legal text changes compliance costs and to estimate whether these costs are low or high. If the costs are likely to be low, the Office will use AI to derive the compliance costs. But if the costs are likely to be high, the estimation will be done manually by humans to help ensure accuracy. This process should allow the Office to focus on high-effort estimates and free up resources for other projects. However, there are still challenges relating to the structure of data scraped, the understandability of German legal texts, data quality, explainability of variables used in the system and matching across data sources.

Source: (Walprecht and Lewerenz, 2024[30]).

AI can also enhance policy assessment and support more informed, iterative decision-making by enabling sophisticated policy experimentation and evaluation. Through simulation of regulatory scenarios of the future, AI allows policymakers to model and predict the potential impacts of different regulatory choices (OECD, 2025[31]). This helps with regulatory impact assessments, as it supports decision-makers understand the consequences of proposed regulations on different sectors and stakeholders. It can help pre-emptively forecast the effectiveness of potential policies by uncovering patterns that may not be immediately apparent through traditional analysis methods.

For example, the University of Dublin led the development of the Innovation Policy Simulation for the Smart Economy (IPSE), a tool that simulates the effect of policy instruments based on regional profiles and sector information. The aim is to better understand the potential drivers of innovation and their impact before policy instruments are rolled out (Nesta, 2024[32]). Similarly, PolicyEngine is a tool that uses digital-ready legislation to inform policy changes and model changes to see how changes would affect governments and citizens. Users can select their country (Canada, United Kingdom, United States and Nigeria are currently available), relevant policy area and policy parameters that are to be changed, and calculate the economic and budgetary impact to see how the changes affect government income, for example. PolicyEngine has been integrating ChatGPT into its system to provide further analysis and explainability to users (Martin, 2023[33]).

Promoting stakeholder engagement in regulatory design

While governments are slowly improving their stakeholder engagement activities for regulatory governance, most OECD countries have significant scope to improve these efforts (OECD, 2023[34]). AI can improve the process efficiency and effectiveness in engaging stakeholders when designing policy. It can do so by using advanced analytics and intelligent user interfaces, contributing to a more inclusive regulatory design process, improving governments responsiveness, and enhancing transparency and trust in government. A growing number of AI chatbots are facilitating public consultations on new or revised regulations by interacting with many stakeholders simultaneously, and synthesising data for governments to further adjust their regulatory proposals. Chatbots can also respond instantly to stakeholder queries, guide them through the consultation process, and simplify complex legislative text to enhance the public's understanding. This makes the process of regulatory consultation more accessible to stakeholders and lowers burdens for regulators when it comes to participatory policy design. The "AI in civic participation and open government" section of this chapter further analyses how AI is being used for stakeholder engagement and public participation.

Enhancing economic regulators' functions

Economic regulators handle diverse functions where AI can be applied, including tariff-setting, authorising activities, handling complaints, mediating disputes, monitoring markets, and conducting inspections and enforcement activities. Beyond core regulatory functions, economic regulators carry out many activities to be effective at their work, such as research to better understand the sectors and markets they oversee.

While some efforts are underway, regulators are in the early stages of using AI, with few exploring or piloting its use.[11] According to a survey by the Body of European Regulators for Electronic Communications (BEREC), "the adoption of AI by the National Regulatory Agencies is still at its infancy" and very few regulators have "undertaken or are planning to conduct studies to explore ways how AI can be adopted" (BEREC, 2023[35]). Nonetheless, the current and potential uses of AI in regulatory approaches reflect a broader trend towards increased use of data and digital technologies. OECD (OECD, 2025[31]; 2020[36]) analysis shows that digital tools help regulators by enabling better data use and informed decision-making. Facing pressure to do more with less, regulators are increasingly turning to AI to make processes more efficient and effective.

Recent discussions among the OECD Network of Economic Regulators (NER)[12] and examples in recent literature reveal that AI is already used in inspections, market monitoring and consumer-facing areas like complaints handling or responding to consumer queries. For example:[13]

- EU e-communications regulators are using AI for radio channel modelling and optimising spectrum sharing, detecting illegal or prohibited content online, conducting customer relations management such as complaint classification, and measuring people's online experiences and platform behaviours at scale. This can be coupled with monitoring the market to ensure that products for sale online comply with product safety rules (Box 5.14) (BEREC, 2023[35]; Faculty, 2021[37]).

- Austria's energy regulator (E-control) is developing an AI application to help consumers understand their energy bills and a chatbot to respond to consumer queries.

- Peru's water regulator (Sunass) (2024[38]) is applying AI in the development of inspection reports. The application automates the generation of reports based on variables recorded by inspectors in tables, significantly simplifying the process and reducing the time spent on report writing. The reports are validated by the specialists to ensure their accuracy. Sunass has also developed a tool that uses geospatial analysis and a classification algorithm to calculate the investment needs and gaps in Peru's water sector.

- The Israeli Capital Market Authority has begun developing an AI-powered tool that aggregates information from websites on insurance and savings. The project integrates advanced financial models, ML and NLP to enable earlier and more accurate identification of risks, anomalies and suspected cases of fraud in the capital markets. Additionally, it provides tools to increase transparency for the general public and investors. The system — currently at the proof-of-concept stage — offers visual tools for training models, monitoring performance, decision-making and generating actionable insights for Authority employees.[14]

- Brazil's National Agency for Land Transportation (ANTT) uses AI in its supervision of transport infrastructure to support an effective delivery of its mandate as an economic regulator in the sector. The Road Information System combines data on aspects including accidents, roadside assistance, possible offenders, toll gates, speed cameras and traffic sensors on 26 concessionaires. The system records 15 000 entries per second with real-time data and combines AI tools with a human interface with a team on a 24/7 basis. The information allows the ANTT to supervise activities in the road transport sector more effectively and supports data-driven regulatory decision-making by the ANTT and other public actors.

Box 5.14. Denmark's SAFE AI tool

Denmark developed an AI tool in 2021 to scrape the internet for dangerous products in 16 European countries, thereby automating a process that previously required manual searches. The tool, called SAFE, was developed by the Danish Safety Technology Authority in co-operation with a private IT company to automate resource-demanding work and improve safety for consumers.

The SAFE tool uses image and text recognition to search the web for dangerous or deficient products, using information from Safety Gate, the European rapid alert system for dangerous products (RAPEX), and the Information and Communication System on Market Surveillance. The tool is continuously trained by using user feedback to improve the accuracy of its findings. SAFE's findings can be used to warn authorities of any non-compliant products for sale in their markets.

SAFE follows on an earlier AI tool by the authority, called AIME, developed for the Danish market in 2020. The authority received funds from the European Commission to develop a similar tool for use throughout the EU, which led to the creation of the SAFE tool.

Source: https://www.sik.dk/en/business/nyheder/new-digital-fine-toothed-comb-denmark-about-make-more-300-million-european-consumers-safer.

Refining validation of risk criteria

Effective regulatory delivery depends on an accurate understanding of risks. Public authorities should base their inspection and enforcement activities on risk criteria. Therefore, it is critical to monitor how risks evolve in real life, enable adaptive responses to possible changes and regularly update risk criteria.

ML techniques can calculate the accuracy of risks assessments and refine parameters (Box 5.15). Recent ML applications are promising in identifying key risk predictors, considerably improving the effectiveness of risk-based targeting (OECD, 2021[28]). As ML algorithms evolve with new data, they need to be updated regularly to remain reliable (Cary Coglianese, 2024[39]). While these tools have considerable potential, regulators need to use discretion when using AI applications to predict compliance to ensure accuracy and avoid biases.

Box 5.15. Risk criteria for agricultural contributions in the Autonomous Province of Trento

In Italy, the Autonomous Province of Trento worked with the OECD to analyse the risk parameters used by the Agricultural Payments Agency (APPAG, Agenzia Provinciale per i Pagamenti), propose revisions and introduce robust ML standard practices. The APPAG pays agricultural contributions related to the management of agricultural land, such as manual mowing or limited use of pesticides. The contributions are proportional to the land's surface area according to the application presented by the farmer. Because supervision resources are limited, the selection of sites to inspect with the highest possible precision is particularly important. The APPAG's risk criteria to plan relevant inspections were validated by applying ML techniques. An algorithm capable of predicting the most at-risk agricultural contribution requests allowed the revision of existing parameters, making them significantly more effective in targeting risks (high risk requests) and identifying non-compliance situations.

Source: (OECD/EU, 2024[40]).

Enhancing risk modelling to improve targeting of inspections

Regulators can use AI to improve risk modelling to better target inspections. Through a risk-based approach, regulators direct their resources towards activities that pose a threat to public goods and, more broadly, to the achievement of desired objectives. Improving inspection efficiency and regulatory enforcement depends on a precise understanding of risks to the public interest (OECD, 2014[41]). Establishing and evaluating risk criteria is necessary to gauge the level of risk posed by private operators and to target enforcement efforts. This better protects the public good, ensuring the efficient use of resources and fostering a trust-based relationship with businesses (OECD, 2018[42]; Blanc, 2018[43]). See Box 5.16 for an example of this approach.

Box 5.16. Tuscany and the OECD develop risk criteria for grant applications

Regulators' application of AI techniques help enhance the understanding of which characteristics of businesses may be effective predictors of risk. This could considerably improve risk-based targeting. Using AI techniques tools and advanced data analytics, regulatory delivery authorities can tailor their oversight strategies, focusing resources on the most critical areas, minimising risks while ensuring that their actions are both efficient and impactful.

Italy's Tuscany region applied a risk-based methodology to documentary checks on funding requests for economic activities operating in the region, particularly for European grants to incentivise innovation. The OECD helped to develop risk parameters guiding the control of these incentive requests. By analysing the relation between characteristics of small and medium-sized enterprises (SMEs). applying for public funding and a suitable potential-risk estimation, predictive ML systems were constructed. Depending on the type of requests, public administrations could use these multiple systems to guide, enhance and accelerate decision-making regarding grants and funding applications.

Such tools and algorithms could serve as a basis for classification of future applications after updating the systems with the relevant data, as the data sets grow over the years. The use of these tools over multiple years could highlight the relapse associated with malicious non-compliance or the improved behaviour of establishments as they become better at preparing their applications.

Source: (OECD/EU, 2024[40]).

AI-enabled risk modelling can also be enhanced through data gathered from social media. Historical inspection results and databases already serve to better target future inspections and preventive strategies. Regulatory delivery authorities have also started to acquire and use data available on social media platforms to identify potential non-compliances. Whereas traditional sources of information rely on authorities and inspectors, social media allow direct access to citizens exposed to risks. This helps further inform risk-based approaches and ensure decisions are result-oriented.

Inspections should be guided by risk assessment, and public complaints often highlight emerging risks. Better inspections begin with informed and cohesive decisions. Complaints need to be used to improve risk-based inspection planning, and should, only in a few cases, lead to impromptu inspections. A risk-based complaint management system is therefore essential to ensure the right balance between proactive inspections — which occur after careful risk-based planning — and reactive inspections — which are unplanned and occur ad hoc in response to complaints deemed serious.

Official complaints are frequently lodged through dedicated apps. However, the public might be reluctant to use these initiatives or simply ignore that this can help track health and safety issues. Regulatory authorities could therefore use social media platforms and ML techniques to support risk analysis and gain

broader and timelier insights than through traditional methods (OECD, 2021[28]). The OECD supported an initiative in Italy's Lazio region to test complaints on social media as a source of the CMS inspired by international initiatives (Box 5.17).

Box 5.17. Enhancing risk criteria through customer complaints in the Lazio region

Regulatory authorities need to use risk assessment to target their enforcement strategies to avoid or address the most likely or most severe potential negative impacts of non-compliance. Citizens' complaints available on the internet could be an essential source of information and play a crucial role in enforcing compliance with existing regulations. By improving their tools for collecting and analysing public feedback, authorities can enhance risk identification and inspection planning.

When citizens do not use the official channels to provide their feedback, including hazard complaints, they are likely to do it through social networks and other websites. Two million TripAdvisor restaurant reviews were collected, and a sample of 5 000 comments was selected. Each review of this sample was manually categorised based on the presence of hygiene issues or food poisoning. This pre-classified dataset was used to train a ML algorithm, an approach based upon a bidirectional long short-term memory (LSTM). This LSTM algorithm was then applied to parse new reviews on the website and classify them, identifying negative reviews (with hygiene issues). Even with the limited pre-classified data used to train the algorithm, the system's performance ranged between 81% and 83%, showcasing its efficiency.

Source: (OECD/EU, 2024[40]).

Improving non-compliance identification

Smarter compliance monitoring and targeting of non-compliance situations should be underpinned by accurate data repository. When inspectors are challenged by the increase of potentially suspect data, ML solutions can help identify non-reliable data and spot non-compliance.

Innovative tools streamline data submission, detect anomalies and analyse compliance pattens (Box 5.18), that enhance oversight efficiency for regulatory delivery authorities.

Box 5.18. Good standing approvals and compliance in Israel

In 2024, the Registrar of Charitable Trust Unit at Israel Corporations Authority's Ministry of Justice introduced an AI-based automated process to its core regulatory approval system for issuing "Good Standing" (proper management) approvals. This approval is vital for over 23 000 non-profit organizations annually, as it serves as a prerequisite for donor tax refunds and eligibility criteria for public benefits and procurement opportunities.

The new automated AI-based system overcame significant optical character recognition (OCR) and object detection challenges to identify compliance with proper management standards while flagging suspicious signs of corruption or improper management. Supervisors can now focus on flagged cases, while average response times have been reduced from 45 days to one hour — delivering a new level of government service.

> The project was one of nine winning initiatives in an AI implementation open call that was announced by the Ministry of Innovation, Science and Technology, in collaboration with the Israel National Digital Agency.
>
> Source: Government of Israel officials, https://www.gov.il/en/pages/most_ai_government_agencies_open_call_winners.

Managing risks and challenges

Associated risks

- Inadequate or skewed data in AI systems
- Lack of transparency and explainability

If AI systems rely upon **inadequate or skewed data**, it could lead to inaccurate or adverse outcomes for some individuals or groups. With regard to PFM, this could result in, for example, adverse regulatory outcomes where some individuals or groups are improperly targeted for enforcement action.

Advanced AI systems often make decisions within a black box, often without even the system operators understanding how it arrived at the decision (Valderrama, Hermosilla and Garrido, 2023[44]; OECD, 2024[29]). This **risk of limited explainability**, necessitates human oversight and evaluation of the AI system and its outputs to ensure transparent and accountable regulatory decision-making. If AI is used for regulatory design or delivery — such as drafting texts, conducting assessments and engaging with stakeholders — careful attention should be given to: the inputs and outputs of such an application; the quality of data used; the accuracy and reliability of AI outputs; the explainability of such outputs; transparency of AI's use in decision-making; and accountability for the associated impacts on regulatory design and delivery.

This risk can make it **difficult to promote accountability and build trust** in governments' ability to use AI to improve rulemaking. Government AI systems should generally be answerable and auditable, which helps to reinforce the OECD AI principle on accountability. As practicable and appropriate, governments should prioritise making AI systems open and transparent to foster public trust and enable external scrutiny and validation. This can include making data public, open-sourcing algorithms and ensuring transparent decision-making processes to boost confidence in AI-assisted decisions. Additionally, clear structures need to be in place to ensure appropriate accountability and oversight mechanisms — considering who is responsible for each element of the AI system's output and who is accountable to the quality or review of outputs across the AI initiative.

Implementation challenges

- Inflexible or outdated legal and regulatory environments
- Lack of high-quality data and the ability to share it
- Skills gaps

While AI offers many opportunities for adaptive regulatory governance, frequent changes to regulatory design can disrupt both business and the public. **Adjustments based on continuous data analysis may lead to a volatile regulatory environment**, making it difficult for businesses to plan long-term strategies and for the public to stay informed about current laws. Further, policymakers and regulators can face **challenges in data access, collection and processing** — which limits the extent to which AI systems will conduct reliable analysis and valid recommendations. For example, Australia has a Data Availability and Transparency Act that provides a legal basis for sharing of Australian Government data. However, in

many instances data-sharing is stifled by a lack of protocols or incompatible systems (Productivity Committee, 2024[45]). Improving data management capabilities is foundational to any use of AI that is assessing and informing the evidence-base for regulatory design or delivery.

As AI in regulatory governance still requires human intervention, a lack of expertise can lead to poor outcomes and misuse of AI. Data collected by the OECD through discussion with its members shows that regulators struggle to match the expertise of technology companies due to the high cost and scarcity of digital skills. To bridge this gap and enhance regulatory effectiveness, **governments need to invest in building digital skills** and promote collaboration with the technology sector, including through partnerships and public procurement to acquire skills and capacities if they are not already present.

Untapped potential and way forward

AI in the design and delivery of regulatory governance is still in its early stages compared to commercial applications. However, there is greater scope for AI to be used in regulatory policy—with a number of key themes emerging from case studies that governments can consider—to advance AI maturity in regulatory design and delivery. With regards to regulatory design:

- AI can be further **leveraged for the design of regulations**, with the extent of its potential application currently under invested. While AI is already being used by governments, many of these applications focus on operational decision-making, compliance measures, general internal processes and delivery of public services or products. Yet these AI applications can be easily adapted to regulatory design. For example, chatbots used in public service delivery or application and grant reviews can be repurposed to gather stakeholder feedback and analyse legislation for regulatory review. AI used for service delivery and tracking market activities can also be used to continuously monitor the impacts of existing regulations, providing real-time feedback and allowing for timely adjustments by regulators.

- **Making legislations more digital ready allows broader AI application** throughout the policy cycle. AI can be used to simplify legislation, make language tech-neutral and improve design for automated case processing. If legislation can be partially, or fully administered digitally, then AI can be used to support delivery, compliance and ex-post review as well as generating data to better inform regulatory policy design and experimentation.

- **Using AI to anticipate future scenarios and risks for more informed regulatory design**. AI can offer valuable forecasting insights, offering governments an ability to see emerging trends and shifts in various industries to proactively plan regulatory responses. For example, AI can forecast how technologies or technological applications grow in the healthcare industry, allowing regulators to test existing frameworks or develop new ones to ensure safety and efficacy before widespread adoption. This can help increase trust in government, as governments take a proactive, rather than reactive, role in citizen protections (OECD, 2024[46]).

As for other policy areas to advance the use of AI for regulatory design, governments need to address legal and governance uncertainties. Many governments may not have the appropriate legal structures and frameworks in place to confidently deploy AI for regulatory governance. This may be due to, for example, ambiguities in legal text about compliance and accountability, which require legal judgement. While countries are taking steps to clarify AI use in government, more mature guidance is needed on the use of AI in regulatory design. Stronger AI governance is crucial, not just for regulatory systems but all government applications to ensure responsible and trusted AI deployment.

Public authorities also need to adapt regulatory delivery mechanisms to provide good protection to public goods and citizens effectively in the context of globalisation and technology, harnessing these changes to achieve better results. Steps forward include ensuring that regulators have the necessary mandates,

powers, functions and accountability mechanisms to use AI; and providing access to high-quality, accurate and secure data to protect against cyber threats. Three main areas of attention lie ahead when deploying AI for regulatory delivery:

- **Empowering regulators with appropriate data collection powers**. To effectively gather and make sense of data from a variety of sources, regulators need sound legislation. Regulating data-driven markets requires an appropriate toolbox, so that regulators can request and receive sufficient information in real time and in relevant shape (OECD, 2020[36]). As AI applications often rely on mass data collection, regulators should be legally empowered to collect, process and publish data when appropriate, while upholding principles of privacy and data protection, and ensuring that legal and institutional conditions support data access and sharing in ways that advance the public interest (OECD, 2021[47]). Regulators need to collaborate with national governments and parliaments to ensure the appropriate legal framework that defines their remit and powers.

- **Enhancing regulators' skills and knowledge on AI**. To unlock the potential of ML techniques and big data analytics, regulators need to boost their AI expertise by recruiting data scientists and cyber risk experts. Attracting and retaining data talent amidst private sector competition is a challenge for regulators and the broader public administration. Further, there should be investment in capacity building for every policymaker and regulatory to boost policymakers' confidence in AI applications. These efforts are at both the national level and international level. For example, new fora for international exchange and co-operation can be developed for officials to share their knowledge and experience — such as the International Network for Digital Regulation Co-operation (INDRC) established to foster discussion between regulators (DRCF, 2023[48]).

- **Ensuring robust data governance and strategies to underpin AI use**. The digital transition offers benefits but also poses new challenges for data governance (see Chapter 4, section on "Creating a strong data foundation", for a detailed discussion on managing, collecting, providing and using data for AI). Robust strategies are essential to mitigate potential risks. Developing a data strategy is a holistic way to address data governance. A poll among members of the OECD *Network of Economic Regulators* (NER) in 2024 found that 55% of respondents were in the process of developing a data strategy, while 29% already had one in operation.[15] These findings highlight the necessity of robust governance and data strategies to support an effective use of data and AI.

Key government processes

Key government processes — such as civil service reform, public procurement, anti-corruption efforts, policy evaluation and civic participation — are vital to building efficient, transparent and accountable institutions. These processes strengthen public trust, improve service delivery and foster evidence-based governance. Governments are using AI to ensure integrity in how taxpayer funds are managed and spent, as well as bolstering public sector talent for the future.

AI in civil service reform

AI has the potential to significantly transform the way civil servants are organised and managed. Like in other policy areas, AI tools could be used to better target and personalise human resources (HR) services, accelerate and improve HR processes to reduce administrative burdens, and boost employee productivity. Some governments are already experimenting with applying AI to discrete HR functions, including recruitment, learning and employee development processes (e.g. training). This kind of application is generally focused on process automation with the objective of ensuring faster and more accurate HR processes. AI's potential in this field could be much greater. Given the right data, AI systems could eventually help better match people to jobs, and predict performance based on job requirements and individuals' characteristics and backgrounds. However, many elements need to be in place for this to work. For example, public administrations would need to have strong data in at least three areas: the characteristics of their workforce, the demands of specific jobs, and indicators of individual performance. Unfortunately, most public administrations lack strong data in all three areas. To unlock this potential, governments will need to invest in more and better HR data, as well as associated skill sets in the human resource management (HRM) function.

Current state of play

AI could be applied to a wide range of strategic HRM processes and activities, finding patterns in civil service data to inform more strategic workforce planning, better target HR policies or even identify needed skills during emergencies and crises. Available data and use cases indicate two areas of civil service reform where some countries were actively experimenting with AI applications: recruitment; and learning and development. These are explored and discussed further in this section.

So far, initiatives have often been fragmented pilots without a strategic approach for systemic adoption. However, governments like France are taking a coherent and strategic approach to explore the benefits of AI in civil service reform, while also recognising the potential negative impacts and setting up safety measures to address them (Box 5.19).

Box 5.19. France's strategy for using AI in government human resource management

French has developed a structured strategy to integrate AI into human resource management (HRM) across its public administration. This strategy focuses on three key areas: AI integration, workforce planning and training for civil servants. It aims to ensure that AI is used responsibly, ethically and effectively, enhances productivity and supports complex decision-making processes.

The strategy first focuses on identifying the right stakeholders, tasks and tools for AI adoption within HR activities. It includes defining clear objectives for AI use, selecting reliable AI tools and establishing methodologies to ensure transparency and accountability in AI-driven processes. The plan also incorporates risk mapping and internal audits to monitor AI's impact and help ensure its ethical use.

The strategy highlights the need for a strategic workforce planning approach specifically tailored to integrating AI in HR processes. The adoption of AI tools to perform tasks traditionally carried out by

---

as supporting candidates — from when they identify a role through to job acceptance — sifting large numbers of applications, scheduling and planning interviews, and matching candidates on reserve lists to other available roles.

Source: https://www.thetimes.com/uk/society/article/inside-hmrc-job-without-speaking-human-ai-tmrljvv2r.

In addition, AI applied to public service recruitment holds great potential to broaden candidate pools and improve candidate screening. Public administrations often struggle to be more proactive at recruitment, to address skills gaps in their administrations and proactively market job opportunities (OECD, 2021[49]). Examples here include one from Canada in Box 5.21. In addition, in Sweden, Upplands-Bro municipal government developed a tailored AI interview robot with a private AI-consulting company to ensure a more accurate recruitment process while enhancing efficiency. While excluding data regarding age, sex, clothing and looks, the AI bot performs blind interviews and evaluates the first-round candidates.[17]

Box 5.21. Increasing representation of visible monitories in Canadian defence leadership

In September 2020, Canada's Department of National Defence (DND) launched an external EX-01 pilot recruitment campaign aimed at increasing the representation of visible minorities at the senior level. This initiative piloted new approaches while embracing inclusivity, innovative methods and technology in order to make fundamental and long-lasting change that helped increase representation and Diversity and Inclusion (D&I) at the DND. Key objectives of the pilot were to:

- identify opportunities for members of visible minority groups;
- introduce novel tools and technology to support barrier and bias-free assessments; and
- assess the pilot process against traditional approaches to identify systemic barriers and biases and identify recommendations for future recruitment processes.

In partnership with various stakeholders, proper guardrails were implemented to ensure privacy and quality assurance. The process achieved key outcomes by facilitating an objective and fair assessment of candidates, while removing biases and barriers at the various stages. As a result, this pilot created career advancement opportunities and improved traditional approaches and processes to achieve better outcomes.

Source: https://www.canada.ca/en/department-national-defence/corporate/reports-publications/ex-01-report-visible-minorities-recruitment-campaign.

Facilitating learning and development

Civil service reform can use generative AI to create learning content, such as learning modules and course material based on source documents and information. Efforts are already taking root around the world. AI can also be used to make personalised recommendations, and pathfinding for learning and professional development through a complex and large amount of information and data (Johnson, Coggburn and Llorens, 2022[50]). Examples in these areas include:

- The Australian Public Service Commission (APSC) recently collaborated with IBM in 2023 to design, structure and deploy a system that generated course content based on documents inserted from users (Box 5.22).
- Spain's National Institute for Public Administration (INAP) is incorporating AI into the organisation, cataloguing and search functions for its digital platforms, learning offerings and broader library.

The resulting "knowledge graph" makes a large portion of the resources easily findable, searchable and shareable with Spanish civil servants, partner countries and the broader public.

- Korea's Ministry of Personnel Management's learning and development platform incorporates AI-enabled functionality to help sort, organise and recommend content. The AI system being implemented is intended to analyse the user's role and learning history to recommend training and material to develop certain skills from the platforms vast catalogue of 1.4 million pieces of content (OECD, 2023[51]).

Box 5.22. Generating learning content in Australia

The Australian Public Service Commission (APSC) ran a six-week pilot project to use AI to design, structure and deploy an online learning course on digital skills for leadership. The pilot system allowed practitioners to "feed" AI a variety of information materials such as articles, books and speech transcripts into a closed system of information that was used to create the course content. The system created a course outline, objectives, modules and content, followed by a quiz. Pilot findings showed that around 60-70% of what the system produced was usable, relevant, well-structured and accurate. In a survey of users, 91% found the pilot output valuable. To increase these figures, the practitioner creating the content could give feedback to the AI system to adjust it.

There were several benefits to this initiative. Firstly, the drafting was very, very quick. Secondly, the information was drawn from a closed system, which eliminated the uncertainty around where information was coming from and if it was correct. The technology highlighted areas of the modules and pointed to where the information was sourced. Thirdly, the system also had programmatic "checkers" built-in to review the material for issues of concern, such as discriminatory language.

While the pilot system could draft course content very quickly, it did not create production-deployment ready content. Issues identified from the pilot were:

- AI can synthesize existing content well but cannot create content that does not exist. For example, it could not write a course on the use of AI in the civil service at that stage as no content existed.
- The reliability of the content still required expert review, and this is typically where bottlenecks in content production already exist.
- The closed system provided more accurate results; however, it is an expensive option if only used for course production.

Post- pilot, the APSC recognises there is value in on-premises, closed AI systems to assist with content creation. It has however referred further investigation of on-premises AI to another Australian government agency, Services Australia. Further observations on Australian government agencies' use of AI can be found on their AI statements.

Source: Information provided to the OECD by the Australian Public Service Commission, https://www.apsacademy.gov.au/news/piloting-generative-ai-address-aps-skills-gap. APSC's AI statement (https://www.apsc.gov.au/initiatives-and-programs/workforce-information/research-analysis-and-publications/state-service/state-service-report-2023-24/fit-future/supporting-safe-and-responsible-use-artificial-intelligence), Services Australia's AI statement (https://www.servicesaustralia.gov.au/automation-and-artificial-intelligence-ai-use).

Evidence of impact

Given the very nascent nature of the applications described above, it is still too early to provide empirical evidence of impact. Early experiments like those discussed above show significant potential to reduce time and effort required to handle large volumes (e.g. number of applicants, amount of existing learning

material), broaden candidate pools, and reduce human error in decision-making. However, in most cases, there is a notable lack of rigorous, empirical evidence demonstrating their effectiveness and impact. Many AI implementations are based on theoretical potential or anecdotal success stories rather than robust, scientific evaluations. Addressing this limitation requires a concerted effort to design and conduct rigorous evaluations of AI applications in government HRM. Future implementations should be grounded in empirical evidence and tailored to the specific needs and constraints of government organisations.

Furthermore, there are many unanswered questions regarding potential negative impacts. For example, AI hiring tools may help to reduce human error, depending on how they are used. Yet they may also limit the autonomy of hiring managers in ways that impact the quality of hire related to culture, fit in team, or other hiring decisions. They may also further limit managers' abilities to build their own teams to achieve the results they need to achieve. The real quality of decisions made through AI systems are difficult to assess since it requires a longer-term view on performance and job fitness. Furthermore, there are limited baselines to measure against. Traditional (i.e. pre-AI) systems and processes face challenges in measuring and assessing this, and there are no agreed standard indicators, especially in systems without objective performance measures as can often be the case in public administrations.

As such, it may take a long time before governments can adequately measure the real longer-term impact of AI-driven decision quality versus those made by traditional systems or humans. It will also be very hard to quantify AI's impact on productivity of HR systems, as very few standard comparable measures and benchmarks exist. The OECD is currently working with a core set of member countries to try to establish these kinds of indicators, which may help to track improvements driven by AI in the future.

Managing risks and challenges

Associated risks

- "Automation bias"
- Inadequate or skewed data in AI systems
- Misuse or questionable use of AI, resulting in surveillance and privacy concerns
- Lack of transparency an explainability

In the domain of civil service reform there is **a documented risk of so-called automation bias,** whereby humans prefer not to second guess the results of automated decision aides, even when they have the ultimate responsibility and accountability to take the final decision. Making best use of AI for civil service reform will require upskilling strategic analytical capabilities in many HR activities and among hiring managers (Broecke, 2023[52]).

Another well-documented challenge relates to avoiding, detecting and addressing **partiality in the AI systems** themselves, particularly if AI is being used to inform decision-making related to job selection and career advancement. The problem is that any organisation's historical data is based on past decisions made by humans, and there is a significant risk that AI systems will hard code these outlooks into their algorithms. Add to this the general lack of good employee data and performance indicators, and it becomes difficult to see how quality tailored advice could be given through such AI systems. Introducing any system in this area will require careful monitoring and evaluation mechanisms to detect and correct for bias (Johnson, Coggburn and Llorens, 2022[50]).

The **data and privacy rights of public servants require particular attention** in a public service context, where values such as merit and fairness guide recruiting, inherently limiting the types of data that can be used in AI systems. For example, some private sector recruitment tools regularly check applicants' social media accounts and use this data to assess candidates. There is little empirical evidence that social media posts or physical features that may be assessed in video interviews have any real bearing on job

performance. This raises both ethical and effectiveness questions about many AI-driven assessment tools currently available on the market (Broecke, 2023[52]).

The use of "algorithmic management" tools — software to automate aspects of management in, for example, the allocation of work schedules, the monitoring of work activities or the setting of worker targets — is increasing significantly, reaching an adoption rate of 90% in US firms and 79% in the European Union (Milanez, Lemmens and Ruggiu, 2025[53]). Up to now, government-specific studies have not been conducted. While some of these tools may help raise productivity when applied effectively, tangible concerns have been raised about existing negative impacts of AI and algorithmic tools on job quality, including work intensification, increased stress and perceived reduction in fairness (OECD, 2023[54]). AI could make jobs less fulfilling by incentivising **new types of surveillance in the workplace** that could harm mental health (APA, 2023[55]), or new forms of hyper-efficient yet exhausting "digital Taylorism" in which work is subject to increased surveillance and regulation (UC Berkeley, 2021[56]). AI task management may also have the potential to erode the autonomy and voice of workers, reducing human insights into how work is managed (Gmyrek, Berg and Bescond, 2023[57]). Many existing HR AI tools are designed to "optimise" workforce management (e.g. monitor, control, reduce autonomy in decision-making and problem solving), going against decades of science that shows how employee empowerment builds engagement, performance and trust. Some research suggests that framing AI as a tool used to support employees, rather than replacing them or limiting their autonomy, is critical for fostering positive perceptions of AI in the workplace (Brougham and Haar, 2017[58]).

The more complex AI systems and predictions become, the less they can be understood and explained. This reduces accountability if employers cannot explain their choices and it impedes the ability of employees to understand how to develop themselves to advance in their careers. Merit based recruitment systems are a bedrock of well-functioning public employment systems, and these require transparency and accountability to function appropriately. Employees and their employers need to clearly understand why appointment decisions are taken, and how the skills and performance of individuals are analysed (Cappelli and Rogovsky, 2023[59]).

Implementation challenges

- Lack of high-quality data and the ability to share it
- Explainability
- Skills gaps

Quality data is essential to implementing AI in civil service reform. Unfortunately, **OECD countries lack large amounts of data in most of relevant areas,** and it is often not standardised systematically across organisations to allow for more rigorous and predictive analysis. Descriptive data is often limited to age, sex, education level and career path. Performance is very hard to assess objectively and consistently across teams and organisations. Job roles are also often categorised broadly. If AI is based on bad or incomplete data, it will make bad predictions.

Implementing AI in civil service management systems requires HR professionals with the right skills and mindsets. While AI technical skills may not be required, HR leaders would need to understand the potential application of AI to their systems and have the right skills to be smart buyers of tools on the market. HR professionals working with AI tools often need analytical capabilities to understand the principles of the tools and their use of data analytics, to interpret and challenge results. While abundant in public administrations, **needed AI skills are often lacking in HR departments.**

Untapped potential and way forward

Strengthened analytics for the present and future

Government organisations could gather a great deal of data and information on their employees that could be analysed to improve performance and employee experience; however, AI is generally under-used in the field of HRM due to several challenges. AI capabilities could allow HR practitioners and management to examine current workforce trends (ageing, skills and performance, compensation) to provide key insights into the main challenges and questions of the day: attractiveness or competitiveness of government as an employer; reskilling and upskilling needs; better targeted learning and development opportunities; or the drivers of employee and team performance and satisfaction.

AI algorithms can use time series data for predictive analytics to identify trends and make predictions on the civil service about the future. While simple regression analysis can be achieved without AI, more sophisticated operations could be developed with more complex modelling and scenario building. For example, organisations could potentially reduce employee turnover by predicting high-risk employees for attrition based on their tenure in a position, their levels of team engagement, and other factors. Through the analysis of big datasets, AI can identify factors and patterns that lead to excessive turnover — which is costly to organisations and a detriment to performance — and allow practitioners to take anticipatory steps for improvement. Aside from turnover, predictive analytics can assist workforce planning by anticipating skills or personnel shortages, predicting top performers for specific kinds of roles, supporting diversity and inclusion, or boosting engagement and well-being.

Currently such applications are extremely nascent in government workforces; the OECD was unable to identify many concrete use cases yet in this area. This is likely due to a variety of challenges, such as those listed above, as well as valid ethical and privacy concerns, which are further detailed below.

Way forward

AI tools can provide leaders with a new opportunity to develop a strategic vision and direction for their public service and the HR activities needed to achieve it. AI has the potential to reshape the workforce and augment its skills in many areas. In the field of HRM, AI can speed up HR processes, better target services, knowledge and recruitment/branding campaigns, and generate valuable insights for HR managers and senior leaders on a range of issues from future skills gaps to hiring effectiveness. AI can be a transformative tool in learning and development, bringing knowledge to public servants while building their essential skills and capabilities. This all depends on a clear vision for what the future public service should look like, backed by resources and capable HR teams. This implies a joined-up strategy for workforce development in which AI has a key role to play.

Enhance transparency and give employees an explanation and the ability to contest automated decisions. Studies suggest that many candidates and employees may perceive automated assessment processes as fairer than those conducted by humans if they understand how they work and why they are being used, so long as they are confident that a human will be accountable for the final decision (Broecke, 2023[52]). This transparency should include the inputs to the decision, ensuring care is taken to obtain informed consent and manage employee privacy issues. Governments should have a clear empirical basis for the input data they choose to use.

Include HR professionals and other employees in the design, implementation and evaluation of HR AI tools. Governments should be especially careful when introducing AI tools that may reduce employee autonomy. In some cases, AI-driven automation could reduce autonomy of workers and de-value their expertise, resulting in lower motivation, engagement and commitment. This is especially true when AI tools are directly applied to optimising their productivity, directing them in how to use their time and monitoring their work activities. This can have unintended adverse consequences; it may increase stress and anxiety,

which could increase work absence and turn-over, thus reducing productivity in the longer run. These consequences can be avoided by including HR professionals and other employees in the design, implementation and evaluation of HR tools.

Upskill and reskill HR professionals for the age of AI. Automation stands to hold many advantages and reduce administrative burden for HR professionals and the employees they support. They can take over many of the repetitive and dull tasks leaving HR professionals to focus on more complex and higher value-adding tasks. Those can include dealing with complex cases, recruiting more specialised workers and developing strategy (OECD, 2024[60]). But this will require upskilling and reskilling within the HR profession in many cases. For some types of employees, governments will need to invest in upskilling in more technical areas, like data science and programming, help HR professionals understand and use AI systems effectively.

If AI systems are used as an input into candidate assessments, governments should have well-qualified humans interpreting the results and making the final decisions. Governments should take care to design the process in a way to minimise "automation bias". This may include conducting traditional assessments first, and then adding AI information after, to provide additional insights on a short list of candidates. In this way, AI can help to audit recruitment practices and improve human decision-making without replacing it. This is essential to ensuring the right level of accountability necessary for merit-driven decision-making in public services.

While incorporating AI into the field of people analytics, governments should pursue the benefits of AI while accounting for costs and risks. They need to also have certain prerequisites in place, such as rigorous, trustworthy data, statistical capabilities for understanding and verifying AI-generated outputs and analysis and their weaknesses, and measures to protect employee privacy and avoid bias.

AI in public procurement

With public procurement representing about 13% of GDP in OECD countries (OECD, 2024[61]), adoption of AI in public procurement is often driven by the need to enhance efficiency and operational decision-making and reduce costs (Friton et al., 2024[62]; Hickok, 2022[63]). It is also used to help address challenges such as workforce constraints. With AI, public procurement can become more dynamic and responsive, capable of meeting the demands of a rapidly changing environment throughout the procurement lifecycle (Hickok, 2022[63]; Gastaldi et al., 2024[64]) (Figure 5.3). This digital transformation also presents an opportunity to fundamentally rethink public procurement and administration, improve connections between public entities and suppliers, and enable more dynamic collaboration (Glas and Kleeman, 2016[65]). Yet realising AI's full potential requires effective implementation, robust data governance and a user-centric approach.

Figure 5.3. Potential use of AI and data analytics throughout the public procurement cycle

	Pre-tendering phase				Tendering phase				Post-award phase		
	Needs assessment and market analysis	Planning and budgeting	Development of specifications and requirements	Choosing the right procedure	Request for proposal / bid	Bid submission	Bid evaluation	Contract award	Contract management	Order and payment	Reporting / performance evaluation
AI	Forecasting, automated data analysis, identification of risks and opportunities.	Optimising budgeting, predicting costs, analysing various budget scenarios.	Automation of specification development by document analysis and information extraction.	Optimised selection process, reduced decision-making time.	Automating the drafting of RFPs by analysing past documents and current requirements.		Automating the evaluation process.		Reduced manual oversight, automated monitoring/ compliance checking.	Automated order processing and payment verification.	Accurate reporting on supplier performance and procurement outcomes.
Big data and analytics	Analysing large datasets for market trends, demand assessment, mapping supplier capabilities.	Data-driven budget decisions, identification of cost-saving opportunities.	Determination of most effective specifications based on past procurements and market analysis.	Analysing previous procurement outcomes to identify the most effective procedure.			Data-driven decision making, identification of patterns and anomalies.				Generating insights into supplier performance, contract compliance, procurement efficiency.

Call for tender — Signing of contract — End of contract

Note: RFP = Request for Proposal
Source: OECD's illustration based on desk research and (EC, 2020[66]).

Current state of play

Public entities are integrating AI and algorithmic decision-making into their processes, using these technologies to improve services, streamline operations and enhance decision-making, as well as strengthen risk management, oversight and accountability. A recent study mapped AI-based functionalities offered by public procurement platforms, revealing that 54% of solutions support pre-tendering and planning, 31% focus on tendering activities, and only 4% address the more operational activities of the supply phase (Guida et al., 2023[67]). Additionally, 11% of solutions, including digital assistants and automation of non-value-added activities, support the entire procurement lifecycle. Common functionalities

include spend analysis, risk management, supply chain finance, supplier scouting and negotiation optimisation.

Streamlining operational tasks

AI can help to classify spending to standardise reporting. Classifying spend data into standard taxonomies in public procurement means organising spending into predefined categories, which helps in tracking and analysing expenditures more effectively. Its importance lies in the ability to better make decisions, enhance transparency and identify opportunities for cost savings. For example, Ukraine's ProZorro e-Procurement system uses a ML solution to predict the correct Common Procurement Vocabulary (CPV) code for products and services (Box 5.23). These codes establish a single classification system for public procurement; aimed at standardising the references used to describe procurement contracts to increase transparency and make it easier for potential suppliers to identify opportunities (EC, 2020[66]).

Box 5.23. Ukraine's ProZorro e-Procurement System

Recognising the challenges raised by lack of fair competition and the risks of corruption due to procurement practices, Ukrainian civil society organisations, expert associations and government collaborated in 2014 to develop ProZorro. This open-source procurement system uses AI and advanced analytics to enhance efficiency, transparency and accountability. A decade after its inception, ProZorro has evolved into a data-driven ecosystem that integrates cutting-edge business intelligence (BI) tools to support evidence-based decision-making across government, oversight bodies, businesses and civil society.

The system's BI ProZorro and ProBI analytics modules enable real-time monitoring and deep analysis of procurement data. The public BI Prozorro module offers 49 dashboards covering all procurement stages, allowing users to assess price trends, medical procurement, buyer performance and risk indicators. The ProBI module provides a report-building tool for advanced users to create customised analytics, facilitating regulatory oversight and strategic procurement planning.

The impact of ProZorro's analytics is substantial. These tools have been widely adopted, with over 30 000 users analysing Ukraine's procurement market annually, covering transactions worth the equivalent of EUR 21 billion in 2023. More than 80 procurement studies have informed regulatory improvements, including EUR 250 million in savings since 2021 from policy changes related to tender corrections. Additionally, AI-driven risk assessments help oversight bodies detect irregularities, while procurement entities use data-driven insights to optimise purchasing strategies. The system's collaborative governance model, integrating government, civil society and private sector stakeholders, promotes continuous innovation and responsiveness to emerging challenges.

Source: https://prozorro.gov.ua.

AI can help governments to enhance processes by simplifying and rapidly streamlining highly rule-driven, end-to-end workflows. Simplifying and streamlining highly rule-driven, end-to-end workflows in public procurement involves reducing complexity and automating tasks, such as with the introduction of RPA or LLMs to enhance efficiency. This can be achieved by integrating digital tools and technologies to minimise errors and accelerate the procurement process. For example, in Chile, ChileCompra has transformed how public procurement is performed and has evolved over time to use novel technologies to support government teams (Box 5.24).

> **Box 5.24. Chile advances public procurement with ChileCompra**
>
> In Chile, public procurement has been transformed through ChileCompra, the country's central purchasing body established in 2000. ChileCompra operates Mercado Público, an electronic platform that centralises and streamlines the procurement of goods and services for public entities. The platform supports framework agreements, which allow multiple suppliers to provide products under standardised terms, fostering inclusiveness and competition. Over time, ChileCompra has become the largest virtual store in Chile, promoting transparency and efficiency in public procurement. Additionally, it has introduced innovative tools like ChileCompra Express, an online marketplace enabling direct purchases from pre-approved suppliers without additional tendering processes.
>
> ChileCompra has continually evolved to address challenges, such as uneven supplier participation and inefficiencies in framework agreements. By 2014, over 850 public entities were using the platform, generating approximately 810 000 purchase orders worth USD 1.8 billion annually. Between 2010 and 2015, the number of transacting suppliers increased by nearly 180%, creating opportunities for small and medium-sized enterprises (SMEs). However, the system faced operational challenges, including a high concentration of revenues among a small number of suppliers and limited second-stage competition in framework agreements. To address these issues, ChileCompra redesigned its framework agreements by standardising product categories and introducing competitive mechanisms that reduced prices by up to 28% compared to market rates.
>
> In recent years, ChileCompra has integrated AI to modernise procurement practices further. It introduced standardised bidding templates for AI and data science projects as part of its Ethical Algorithms initiative in collaboration with Universidad Adolfo Ibáñez and BID Lab. These templates include requirements for transparency, privacy, non-discrimination and explainability to ensure responsible AI use in government contracts. Additionally, ChileCompra's Public Contracting Observatory uses AI tools such as LLMs to analyse procurement data for irregularities and improve compliance monitoring. These advancements have enabled more efficient oversight while promoting ethical standards in public procurement.
>
> Source: https://www.chilecompra.cl.

AI can help streamline efforts to manage public procurement legal and regulatory frameworks. For example, in Romania, the National Public Procurement Agency (ANAP) developed a tool to improve its ability to screen new legislation, including retrieving documents in real time from public institutions' websites and converting scanned documents into searchable text (The World Bank, 2023[68]).

Enhancing procurer-supplier relations and public servant capacities

AI provide real-time communication and support through chatbots, which can answer queries, provide updates and facilitate smoother interactions. This technology helps streamline processes, reduce response times and improve overall satisfaction for both parties. For instance, in the United States, the El Paso City Council Purchasing and Strategic Sourcing department (PSS) integrated a chatbot solution, called Ask Laura, on its webpage. Ask Laura uses open-source algorithms to interpret and deal with questions and get information about potential suppliers based on business profiles and the questions asked (Collins, 2020[69]).

AI tools can facilitate real-time support and guidance for procurement professionals and suppliers, streamlining communication, and expediting query resolution. This can help enhance collaboration by offering timely advice, auto-populating forms, and creating personalised dashboards. For example, in the

United States, the North Carolina Department of Information Technology (NCDIT) has introduced an AI-powered chatbot to assist state agency staff with IT procurement processes. Available 24/7, the chatbot provides instant answers to common queries, such as accessing procurement forms, submitting exception requests, and understanding procurement timelines, reducing wait times and enhancing efficiency (NCDIT, 2024[70]).

Improving risk management, oversight and accountability

Automated compliance checks, fraud detection algorithms and anomaly detection mechanisms bolster accountability by flagging irregularities and deviations from established procurement protocols. AI functionalities use ML techniques to automatically identify errors and fraud, and to manage risk efficiently and effectively (Guida et al., 2023[67]).

AI can also be used to identify integrity breaches in procurement. Both administrations and citizens are investigating AI's potential this regard. As one relevant example, using publicly available procurement auction data, researchers have been able to use AI algorithms to detect collusion with an accuracy rate of 81-95%. Once the algorithms are trained, they can be automatically updated with the latest auctions, and with little effort on the user's part to supervise their outcomes (Garcia Rodriguez et al., 2022[71]). In addition, in Hungary, a study analysed 119 000 government tenders from 2011-2020 to identify subtle, text-based strategies used by corrupt actors to favour specific bidders. Using ML approaches like Random Forests, the study found that text data improved system accuracy from 77% to 82%, demonstrating the potential of text mining to uncover corrupt behaviours and enhance anti-corruption policies (Katona and Fazekas, 2024[72]). Another relevant example is from Spain, where researchers developed an AI system to provide an "early warning system" predicting public corruption. The tool uses data on economic and political factors — such as economic growth and the length of a political party's time in power — along with data on corruption cases to predict the risk of corruption in Spanish provinces (López-Iturriaga and Sanz, 2017[73]). Anti-corruption beyond public procurement is discussed below in the section on "AI in fighting corruption and promoting public integrity".

Beyond identifying anomalies in existing data, AI's predictive capacities can flag potential risks and irregularities and optimise public procurement processes. In Brazil, for example, the Comptroller General's Office developed Alice, a tool that uses AI to detect possible instances of fraud, enabling real-time risk management and oversight (Box 5.25).

Box 5.25. Brazil's AI-powered procurement oversight with Alice

In Brazil, public procurement accounts for a significant portion of government spending, making it a critical area for ensuring efficiency and transparency. To address vulnerabilities such as fraud, inefficiencies and errors in the procurement process, the Comptroller General's Office (CGU) developed Alice, an AI-powered system designed to analyse bids, contracts and public notices. Alice uses artificial intelligence and RPA to continuously monitor procurement activities across federal agencies, identifying risks and irregularities in real time. By automating these processes, Alice enables large-scale auditing and supports public officials in making informed decisions to improve oversight.

Since its implementation, Alice has delivered remarkable results. In 2023 alone, it analysed nearly 191 000 acquisitions and triggered 203 audits involving contracts worth EUR 4.15 billion (equivalent). Between 2019 and 2022, its alerts led to the suspension or cancellation of bids totalling around EUR 1.5 billion (equivalent). Additionally, the system has significantly accelerated audit processes, reducing the average time required from 400 days to just eight days. Alice leverages tailored natural language

---

processing (NLP) algorithms capable of handling the unique complexities of Brazilian procurement data, further enhancing its effectiveness in identifying risks across approximately 40 predefined typologies.

By combining advanced technology with robust institutional frameworks, Brazil improved oversight, achieved substantial financial savings and enhanced accountability in public spending.

Source: https://oecd-opsi.org/innovations/robot-alice-bid-contract-and-notice-analyser.

Oversight bodies including supreme audit institutions are key actors ensuring the legality, efficiency and integrity of procurement processes. (OECD, 2015[74]). Every year, Portugal's court of audit (Tribunal de Contas, or TdC) conducts a significant number of reviews and audits related to public procurement processes (before, during and after), requiring extensive human and financial resources. TdC is working with the OECD to develop stronger control capabilities and more efficient allocation of resources. This includes the development and testing of data pipelines and ML systems to identify red flags, focusing on risks of irregularities in public procurement processes (OECD, 2024[75]).

Empowering external actors and strengthening trust in government

AI can also empower external actors, such as citizens and civil society organisations, to conduct third-party assessments of public procurement programmes and government spending (Santiso, 2022[76]). Governments have developed platforms that enable stakeholders to access open public procurement data, facilitating a transparent exchange of information (Attard et al., 2015[77]). These portals publish both structured and unstructured data covering various phases of the contracting processes. To enhance the presentation of this data, some government agencies have integrated AI-driven dashboards into these platforms that display statistics and indicators relevant to the contracting processes (Ansari, Barati and Martin, 2022[78]). Countries including Colombia (2024[79]), Chile (2024[80]) and Mexico (2024[81]) have adopted such initiatives.

AI technologies can enable stakeholders to monitor procurement activities and manage risks with unprecedented granularity by providing real-time access to procurement data, audit trails and performance metrics. Moreover, AI-powered transparency initiatives can promote greater public trust and confidence in government procurement practices, fostering a culture of integrity and accountability. AI tools can help governments to anticipate demand, identify potential risks and optimise their procurement processes. In Brazil, for example, the Labcontas project (GLOBO, 2018[82]), which brought together 96 databases with information relevant to the work of the Brazilian Federal Court of Audit (TCU), has enabled automated checks of public tenders posing a risk of potential corruption (EC, 2020[66]).

Managing risks and challenges

A variety of potential risks and challenges are areas of concern for the use of AI in government (Andersson, Arbin and Rosenqvist, 2025[83]; Shark, 2024[84]). The issues most seen in OECD work and the evaluation of AI used cases are discussed below.

Associated risks

- Inadequate or skewed data in AI systems
- Lack of transparency and explainability

AI systems with **inadequate or skewed training data can present concerns**, as systems used to evaluate bids might favour certain bidders due to skewed training data, leading to unfair procurement

decisions. Moreover, the ability of public procurement teams to understand the accurate functioning of algorithmic systems is constrained by human biases in perception, multi-layered complexities and other factors (Hickok, 2022[63]). These types of AI systems can have harmful outcomes that might impact a larger portion of society. In procurement, a human public employee reviews bids and decides one by one, whereas an AI system can review many bids and make decisions in a matter of seconds and minutes. Therefore, if bias exists in the AI system, the speed and scale of harm will also exceed that of a human review (Hickok, 2022[63]). To prevent these outcomes, procurers need to ensure that the AI has been trained on representative datasets. In addition, the AI systems should be designed with fairness in mind, considering factors beyond cost and efficiency. Finally, AI performance should be monitored in real-world scenarios to detect and address any emerging harms.

Concerns also arise regarding AI's **lack of algorithmic transparency.** When an AI system is deployed within the public administration, governments should make and honour commitments to principles of fairness, accountability and transparency. The system should be sufficiently explainable; the contracting authority should obtain sufficient information on how the system works and the data it has been trained on to derive its conclusions. Otherwise, public actors will have embedded systems without an independent capability to maintain and monitor these systems. Without these capabilities, alternative oversight and accountability mechanisms will also not be available due to the initial lack of transparency or subcontracting arrangements (Hickok, 2022[63]). In the United Kingdom, the Office for AI (OAI) and the Government Digital Service (GDS) produced a guidance in partnership with The Alan Turing Institute to safeguard public trust in the use of AI in procurement through the use of the FAST Track Principles: fairness, accountability, sustainability and transparency (GOV.UK, 2019[85]).

Implementation challenges

- Inflexible or outdated legal and regulatory environments
- Lack of high-quality data and the ability to share it
- Skills gaps
- Risk aversion
- Data and vendor lock-in

Many jurisdictions **lack regulations and formal guidance on AI use**, leading to legal ambiguities and potential challenges from unsuccessful bidders questioning the fairness of the process. Given these regulatory gaps, there is a growing need for regulatory frameworks and guidelines to ensure clear guidance for AI in public procurement, promoting transparency and fairness, reducing legal ambiguities and minimising challenges from unsuccessful bidders. Box 5.26 illustrates how some governments are coming together to overcome this challenge.

Box 5.26. GovAI Coalition for responsible AI procurement and deployment in the United States

The GovAI Coalition is a multi-agency initiative dedicated to promoting the responsible and ethical use of AI in government. Founded in 2023 by the City of San José, the coalition has since expanded to include local, state and federal agencies across the United States. It serves as a platform for cross-agency collaboration, knowledge sharing and AI governance best practices, helping governments simultaneously promote innovation while ensuring accountability.

GovAI Coalition members have collaborated to create a suite of public procurement templates and knowledge-sharing tools that any public agency can use to jumpstart its own AI governance programme. Among these resources, the GovAI Coalition developed the AI Contract Hub, launched in February 2025 in partnership with Pavilion. This platform streamlines AI procurement by offering a shared

---

repository of contract templates, cooperative agreements and best practices. The hub aims to reduce procurement costs and timelines, improve contract transparency and expand access to AI vendors. As AI adoption in government grows — reaching USD 3.3 billion in federal AI-related contracts in 2022 — the GovAI Coalition helps ensure that public agencies have the tools to procure AI solutions efficiently while upholding public values.

Source: https://www.sanjoseca.gov/your-government/departments-offices/information-technology/ai-reviews-algorithm-register/govai-coalition.

AI can significantly enhance the entire public procurement cycle. However, its impact is limited without **standardised and accessible data, which** necessitates a coherent, government-wide data governance strategy. Moreover, government agencies often implement various AI systems without unified standards, resulting in incompatible systems, fragmented data and inefficiencies in aggregating procurement information. Restrictions on data sharing contribute to these challenges and make it challenging to get the most out of AI systems by broadening the scope of system training and analysis (Andersson, Arbin and Rosenqvist, 2025[83]). Although OECD countries have made considerable progress, data containing all relevant public procurement information is still largely unavailable, and very little exists as open data (defined as data reusable in an accessible format) in most evaluated countries (da Rosa, 2023[86]).

Digital skills gaps — as well as a lack of understanding of AI's potential — are relevant hurdles to the successful deployment of AI in procurement processes (Guida et al., 2023[67]). The power of AI alone is not enough for its successful adoption of advanced procurement platforms. Data management, cultural change and skills development are fundamental (Handfield, Jeong and Choi, 2019[87]). If such gaps and limited understanding persist, public entities may struggle to effectively implement and manage AI systems or mitigate their risks, leading to inefficient and potentially improper procurement processes.

Another interesting finding is that **procurement managers are mostly sceptical of AI,** believing that the typical skills of the human buyer are strictly related to negotiation and that this knowledge, often tacit and not formalised, cannot be transferred to autonomous agents or systems (Guida et al., 2023[67]). Consequently, they feel that this knowledge cannot be effectively transferred to autonomous agents or systems. This risk aversion highlights the importance of digital readiness and the level of digital skills possessed by public officers, as these factors are crucial for the successful integration of AI in procurement processes.

Poorly designed or restrictive data licensing agreements can create **data lock-in**, preventing the contracting authority from sharing the necessary data with the AI developer, thus limiting the AI system's effectiveness. And there is a risk of **vendor lock-in**, which makes the contracting authority heavily reliant on the AI vendor's proprietary technology and data formats.

Untapped potential and way forward

OECD findings and external research identify little dedicated research on AI for public procurement, and a fairly low level of AI maturity in public procurement entities (Andersson, Arbin and Rosenqvist, 2025[83]). To help public procurement organisations assess their own maturity and identify factors needed for growth, the IBM Centre for The Business of Government has developed an AI maturity model for public procurement that may be a useful reference (2023[88]).

If successfully adopted in the field, AI's potential in public procurement includes automated supplier evaluation, predictive systems to anticipate events — like product shortages and optimal timing for securing best pricing — detection of potential influences from broader economic and geopolitical issues, and the creation of smart bidding platforms that automatically match procurement needs with the most

appropriate bidders (Shark, 2024[84]). LLMs could be used promote integrity in public spending. For instance, models such as those that power ChatGPT can support public procurement officials in analysing large amounts of data on a company and potential contractor to screen for fraud or corruption risks (Ugale and Hall, 2024[89]).

AI could also make an impact in setting requirements and specifications needed by a purchasing official, assessing bids and ensuring fair and reasonable pricing, optimising supplier selection and ensuring regulatory compliance (IBM, 2023[88]). AI could help public procurement officials to determine specifications for a purchase by presenting information about previous procurement exercises, or by dynamically identifying and presenting relevant products services on the market (IBM, 2023[88]). Improving market knowledge through AI-gathered and synthesised content could also assist procurement officials in identifying reasonable and fair prices for various products and services to help serve as a baseline for direct purchase or in considering competitive bids. Moreover, AI could help to synthesise suppliers' information for procurement decision-making. Information about suppliers can be analysed using NLP methods from a variety of sources, such as business profiles, financial data and internet reviews (Burger, Nitsche and Arlinghaus, 2023[90]).

Other potential applications can be found in the private sector's use of AI tools. Private procurement has been the subject of significantly more research, even though companies are slow to adopt AI relative to other business functions (Andersson, Arbin and Rosenqvist, 2025[83]). These approaches may serve as inspiration for public procurement offices, which may not be able to directly replicate private sector solutions due to legal and regulatory frameworks. Examples include:

- Walmart has implemented Pactum AI, a negotiation chatbot for suppliers providing "goods not for resale", which aimed to enhance payment terms, secure discounts and offer flexible contract termination notices.[18]

- AutogenAI has developed an AI tool designed to expedite the bid-writing process for procurement. This tool assists businesses in crafting proposals more efficiently, reducing the time and effort required in responding to procurement opportunities.[19]

- Sievo, a procurement analytics company, offers a platform to enhance procurement processes that uses AI to analyse spend data, forecast demand and optimise supplier selection, thereby improving decision-making and operational efficiency.[20]

- DocuSign developed an AI-powered contract management tool that uses NLP to scan and interpret legal documents, identifying cost-saving opportunities and ensuring compliance.[21]

Contracting authorities in government need to take steps to ensure they are taking informed actions and decisions when incorporating AI into their procurement processes. They should focus on reducing both risks and risk aversion, improving skills and capacity, encouraging procurement officials to engage in dialogue with suppliers, and enhancing data collection and monitoring of results. The OECD Recommendation on Public Procurement (2015[74]) and the OECD AI Principles (2024[91]) — among other OECD and international standards — help contracting authorities navigate their efforts to implement trustworthy AI in their respective processes. Without an informed and trustworthy approach, public actors could fail to seize the benefits of AI systems. They could also end up with highly integrated yet flawed systems without an independent capability to maintain these systems, or skills to monitor their performance (Hickok, 2022[63]).

The success of AI in procurement processes hinges on both the effective implementation by the buying entity and the commitment of key stakeholders. A user-focused approach is essential for digital transformation in public procurement. Success also depends on robust data governance and infrastructure, including the standardisation, sharing and use of procurement data, as well as modern computing systems needed to efficiently host and transfer data and run AI systems.

Governance mechanisms and accountability structures remain essential. Policymakers and intergovernmental organisations are establishing regulations to govern AI use in public procurement. To ensure solid governance and accountability, AI-specific procurement obligations and documentation should apply equally to both external and in-house development (Heikkila, 2022[92]).

Governments should also invest in capacity development programmes, training initiatives and knowledge-sharing platforms to support AI adoption in public procurement. By investing in capacity building and training programmes, procurement professionals can be equipped with the necessary skills and knowledge to effectively use AI technologies.

Furthermore, collaboration between public and private sectors, academia, civil society and, in some cases, the public is essential for encouraging innovation and disseminating best practices in AI-enabled procurement systems. By promoting knowledge sharing and collaboration, governments can accelerate the adoption of AI technologies and maximise the potential benefits for society. This collaboration also helps to determine whether AI is the best solution for a given challenge relative to other approaches or technologies – an important but often overlooked step (Hickok, 2022[63]).

Public purchasing bodies should prioritise the ongoing evaluation and iteration of AI systems used in public procurement, such as monitoring the performance of AI algorithms. They should assess their impact on procurement outcomes and people and solicit feedback from stakeholders.

AI in fighting corruption and promoting public integrity

Integrity actors — such as anti-corruption agencies (ACAs), supreme audit institutions (SAIs), internal audit and other oversight bodies — are reaping the benefits of adopting AI to enhance their operations, audits and investigations. As data availability intensifies and digital complexity increases, these institutions are forced to adapt to remain relevant. Faced with the increased use of technology by all actors, including those meaning to evade scrutiny and oversight, it is a matter of legitimacy for integrity institutions to keep up with the digital evolution.

AI has been a particular focus for integrity actors in recent years. This is due to rapid technical improvements making these technologies more available and affordable, as well as to the possibilities afforded by AI to process all types of data efficiently, uncover complex patterns and relationships, enhance analytical accuracy and precision, and generally augment current integrity activities (OECD, 2024[93]). As in other policy areas, AI is mostly valuable for the analysis of vast and complex sets of data, where computers can be more accurate or faster than humans.

Current state of play

Across integrity actors, AI has numerous potential applications, which are currently being used, tested or are anticipated to emerge in the coming years. The primary activities have been categorised into three categories based on their potential impact: detecting fraudulent activity; achieving efficiency gains in knowledge management and synthesis; and enhancing predictive analytics and forecasting.

Detecting fraudulent activities

One of the earliest and still most profound applications of AI by integrity actors lies in detecting fraudulent activities, transforming traditional analytical methods into sophisticated, data-driven approaches. AI algorithms excel at applying statistical techniques to identify outliers, patterns, transactions and behaviours that deviate from established norms that warrant further human investigation (OECD, 2024[93]; Taşdöken, 2024[94]). These tools are able to analyse numerical data across entire populations, a significant shift from traditional sample-based analytics, which enables the identification of misallocations or misreporting of funds, among other irregularities (Köbis, Starke and Rahwan, 2021[95]) (Box 5.27). In other examples:

- The OECD recently worked with the Spanish General Comptroller of the State Administration (IGAE) on an initiative proof of concept that employed advanced analytics and ML to detect anomalies and patterns indicative of potential corruption or fraud risks (OECD, 2021[96]).
- The European Court of Auditors (ECA), as part of an audit on EU lobbying activities, has used ML algorithms to check the transparency register of the European Commission and identify outliers with the objective of isolating a more interesting, risk-based sample for analysis by auditors (ECA, 2024[97]).
- In Brazil, the Alice tool analyses daily purchasing and procurement processes to uncover risk areas and inconsistencies. Unusual patterns trigger an alert that suspends the purchase, which is flagged for further inquiry (Box 5.25).

- In the United Kingdom, the Department of Work and Pensions is using AI to identify patterns in claims that could suggest fraud or error so that these claims can be reviewed by relevant teams within the department (Box 5.28).

Box 5.27. Europe's DATACROS tool for detecting corporate ownership anomalies

In the European Union, Project DATACROS has developed an innovative AI-powered tool to detect anomalies in corporate ownership structures that may indicate risks of corruption, money laundering, and other financial crimes. Funded by the EU Internal Security Fund - Police and coordinated by Transcrime at Università Cattolica del Sacro Cuore, the project addresses the critical need for advanced analytical tools in public oversight and integrity efforts.

The DATACROS prototype tool consists of two main components: the Restricted Area for authorised users such as law enforcement and anti-corruption agencies; and the Public Area, accessible to all citizens. The Restricted Area employs AI algorithms to analyse the ownership structures of over 70 million companies across 44 European countries, identifying red flags and hidden patterns that may indicate illicit activities. It incorporates cross-border data and leverages extensive knowledge of criminal schemes, while adhering to privacy and data protection regulations.

Testing activity of the DATACROS tool in 2021 has shown promising results. The predictive systems correctly identified 83% of companies targeted by sanctions and 88% of companies with sanctioned owners. Partner agencies reported high satisfaction (4.3/5) and likelihood of future adoption (4.3/5). By providing both investigative capabilities for authorities and transparent aggregate data for public oversight, DATACROS represents a notable example of how using AI can help enhance public integrity and combat financial crime.

Source: https://www.transcrime.it/wp-content/uploads/2021/10/Datacros_report.pdf.

Box 5.28. Using AI to tackle benefit fraud in the United Kingdom

The United Kingdom's Department for Work & Pensions (DWP) has been using AI systems since 2021-2022 to analyse historical data and identify patterns linked to fraudulent Universal Credit claims. It is investing GBP 70 million in advanced analytics between 2022 and 2025 to reduce fraud and error in benefit claims, with projected savings of GBP 1.6 billion by 2030. The systems adopted assess risk based on inconsistencies in reported income, unusual claim frequencies and mismatches with external records. High-risk claims are flagged for manual review, where payments may be paused until further verification is completed. Key areas of focus include self-employment claims, housing benefit fraud and undeclared capital, where fraud detection has historically been more challenging.

While these systems improve efficiency, DWP recognises that ensuring fairness and transparency remains a challenge. To mitigate bias, it conducts fairness analysis before deployment and continuously monitors flagged claims. However, gaps in demographic data make it difficult to fully assess whether certain groups are disproportionately affected. Caseworkers retain full decision-making authority and are not informed of why a claim was flagged to prevent confirmation bias. Random claim selection is also used to test the accuracy of the system and avoid overreliance on automation.

To strengthen public confidence, DWP has committed to publishing annual reports on the impact of AI in fraud detection, particularly concerning vulnerable groups. Efforts are underway to refine fairness analysis by improving data collection and adjusting risk-scoring methods where needed.

Source: (UK National Audit Office, 2023[98]).

Network analysis is another valuable technique for detecting potential fraud or similarly risky scenarios. It involves examining the interactions among entities within a system – such as public officials, contractors and suppliers – to unearth and visually represent potentially hidden relationships. This method is especially useful in identifying conflicts of interest, as well as in detecting corruption activities that were previously hard or impossible to detect (due to the volume and complexity of the data), providing crucial insights that help maintain the integrity of public operations. For instance, in procurement, pattern recognition can help identify sequences of transactions that may indicate collusion or kickbacks.[22] Anti-corruption in this area is further discussed above, in the section on "AI in public procurement".

Improving knowledge management and synthesis

For integrity actors, who dedicate a large portion of their workday to reading, processing and extracting meaning from text-based documentation, AI can assist knowledge management and text analysis. This can include, for instance, extracting meaning and automating fact retrieval from documents.

Within the realm of NLP, named entity recognition allows the rapid detection and classification of essential information in texts and other unstructured data formats — such as names of individuals, places, dates and organisations — at a speed and scale that would have been impossible to retrieve manually. Named entity recognition is used in chatbots, sentiment analysis, information extraction and fraud detection. Brazil, for example, integrated it into the Federal Court of Accounts (TCU) (Odilla, 2023[99]). Machine translation is also facilitating access to better textual information more quickly, thereby enhancing the effectiveness of audits and investigations.

The rapid development of GenAI is also being used to improve efficiency. LLMs can be of great assistance by allowing the reading, interpretation, summarisation and categorisation of vast quantities of textual information with a high degree of accuracy and in a fraction of the time of a human. Integrity actors are expanding their use of LLMs for internal and public-facing applications, making not only their work more efficient but also the public more engaged and aware of their efforts. Brazil's creation of ChatTCU is one example (Box 5.29). Others include:

- Integrity actors in Finland, France, Greece, Lithuania, the United Kingdom and at EU institutions are using LLMs to support in the drafting of documents, analysing spreadsheets, summarising texts, processing of corruption reports, analysing case law or jurisprudence, and for discovery in performance audit (Ugale and Hall, 2024[89]).

- The Netherland's Court of Audit is piloting a GenAI system to allow citizens and other stakeholders to roam through public reports and to find answers and sources to their questions on public audit work.[23]

- The ECA is piloting a project with three other SAIs (Netherlands, Germany and Sweden) to enhance interinstitutional relations and knowledge sharing by using LLMs to parse through existing national audit reports (across all European languages), automate translation and create English-language factsheets and guided summarisation.[24]

<div style="border: 1px solid;">

Box 5.29. Brazil enhances audit processes through generative AI with ChatTCU

The Brazilian Federal Court of Accounts (TCU) has pioneered the use of generative AI in audit and oversight activities through ChatTCU, an AI-powered assistant launched in February 2023. Developed as an institutional initiative, ChatTCU integrates with TCU's internal systems, providing auditors with real-time access to case summaries, regulatory guidance and administrative support. As of December 2023, the tool had over 1 400 users, demonstrating its rapid adoption across the institution.

ChatTCU, currently operating on GPT-4 32k, enhances efficiency by allowing auditors to retrieve and synthesise information quickly, reducing manual workload while maintaining accuracy and security. Hosted on Microsoft Azure, the platform ensures data protection and prevents unauthorised sharing of sensitive information with external AI providers. Future developments include deeper system integration and workflow automation, further embedding AI into audit processes.

Beyond its internal application, ChatTCU is now part of an international technology transfer effort. In October 2024, TCU signed an agreement with the Supreme Audit Institution (SAI) of Honduras, making it the first international recipient of ChatTCU's source code. This partnership, established through OLACEFS, reflects a broader strategy to strengthen AI adoption in audit institutions across Latin America and the Caribbean. Since its launch, 33 Brazilian institutions have also begun adapting ChatTCU for their specific needs, highlighting its scalability and adaptability.

Source: (Ugale and Hall, 2024[89]).

</div>

With GenAI just entering popular discourse in late 2022, integrity actors have had little time to comprehend the opportunity generative AI presents for their work, let alone to fully integrate it into activities. When the OECD (2024[89]) surveyed integrity actors last year, respondents generally described their organisations as being in the early stages of maturity concerning the use of generative AI and LLMs, if at all, though it is becoming a primary area of focus.

Preventive and predictive analytics, forecasting, foresight

Predictive AI systems can help anticipate potential corruption and fraud risks and other issues, enabling the prioritisation of cases for further human examination. While many integrity bodies lack the mandate for ex ante actions, identifying rising risks early can help mitigate them more effectively. Although ex ante actions are often out of the mandate of many integrity bodies, anticipating a rise in risks associated with corruption and fraud before they manifest can help mitigate these risks. For example, in Colombia, VigIA was developed by the Universidad del Rosario to analyse contracts from Bogotá's Mayor's Office and flag those with a high risk of corruption, anticipating potential illegal conduct. The tool is now adopted by the City of Bogotá (Salazar, Pérez and Gallego, 2024[100]).

Other researchers demonstrated how AI and data science techniques can be adopted to predict corruption. For instance, focusing on Brazil, researchers demonstrated that, in comparison to random audits, an AI system using budget data as predictors could detect almost twice as many corrupt municipalities for the same audit rate (Ash, Galletta and Giommoni, 2020[101]). Focusing on Italy, researchers used police archives and applied AI algorithms to predict white collar crimes demonstrating how algorithmic forecasts could support anti-corruption policy targeting (de Blasio, D'Ignazio and Letta, 2022[102]). In Spain, a research team also developed an AI system that serves as an early warning system to predict public corruption based on economic and political factors across Spanish provinces (López-Iturriaga and Sanz, 2017[73]). The transfer of these approaches from research to governments is still limited but growing. For

instance, in Lithuania, the Special Investigation Service (STT) is developing an AI-based tool to assist corruption prevention officers with their legislative analyses (Box 5.30).

Box 5.30. Corruption prevention through the analysis of legal drafts in Lithuania

The Special Investigation Service of the Republic of Lithuania (STT), the country's main anti-corruption body, is currently developing and testing a tool to help corruption prevention officers evaluate both enacted and draft legal acts for anti-corruption purposes.

The tool uses LLMs specifically trained on corruption risk data and risk assessment methodologies. Through the identification of potential legal corruption risk factors (e.g. loopholes, insufficient or weak procedures and measures), it helps officers rapidly and efficiently analyse large volumes of legislation in order to identify potential shortcomings or vulnerabilities with the legal acts.

By using this tool, STT seeks to strengthen legislative safeguards and ensure that when a draft legal act is adopted, the possible consequences of its implementation are carefully considered.

Source: Ongoing work with the OECD and STT (report forthcoming).

Evidence of impact

AI holds much promise for integrity actors, yet its empirical impact is still difficult to measure at this stage. Anecdotal evidence shows that AI enhances operational efficiency by saving time, allowing smart automations and increasing the accuracy of analyses. Use cases in which NLP, including LLMs, summarise, translate or extract information from text-based sources are among those where AI's added value for integrity-related activities is the easiest to see and assess. However, quantifying the broader impact of AI in the field is complex. First, maturity in the field is generally low; many AI initiatives by integrity bodies are still at the exploratory phase (proof of concepts, pilot projects) and have not yet reached scale.

Additionally, assessing the current state of AI in this field is hindered by a lack of transparency and sharing of information and lessons learned between institutions and across borders. Information remains highly fragmented, and the sharing of use cases, best practices and knowledge about AI tools in integrity remains challenging. The challenges stem from a combination of factors, including trust issues around confidentiality and sensitivity of oversight activities and investigations, and cultural differences as well as varying levels of institutional maturity in adopting AI. Additionally, the rapid pace of technological evolution, coupled with limited resources in terms of both time and expertise, further complicates efforts to share information and learn from diverse use cases. This lack of openness hinders the ability to evaluate and understand the effectiveness of AI applications and learn from different institutions and contexts.

Managing risks and challenges

Associated risks

- Inadequate or skewed data in AI systems
- Lack of transparency and explainability
- Misuse or questionable use of AI

When deploying AI in public integrity and anti-corruption efforts, it is crucial to address the **potential errors that can be embedded in AI systems, such as those caused by inadequate or skewed data** (Köbis,

Starke and Rahwan, 2022[103]). This is particularly critical in areas like fraud detection and conflict of interest analysis, where fairness and accuracy are paramount. Efforts continue to identify objective indicators that could serve as proxies for corruption, a phenomenon that – by definition – is hidden. While the most practical, the reliance on historical data (past cases or decisions) to train algorithms can perpetuate errors. This could result in skewed insights and limit the potential to uncover new and evolving dynamics of corrupt behaviour, potentially leading to unfair targeting or oversight practices, undermining the credibility of integrity efforts. To counter bias, AI systems should be developed with robust checks and balances, including rigorous testing, continuous monitoring and feedback mechanisms that allow for the identification and rectification of unintended errors and resulting harms.

The opaque nature of many advanced AI systems also poses significant challenges for integrity actors. If the public perceives government actions and decisions as opaque and perceive that they are relying on **"black box" algorithms — where AI-generated outcomes lack explainability and transparency** — citizens are likely to resist or distrust the use of AI (OECD, 2024[93]). In the audit profession in particular, its use to generate insights to be used as audit evidence is still a contended issue (Mpofu, 2023[104]). It raises question about their reasonableness, reliability and compliance with professional and regulatory standards (such as International Standards on Auditing, or ISAs) that have not yet been updated to account for recent technological changes.

The integration of AI technologies in integrity bodies raises significant concerns about responsibility, ethics and data protection, especially regarding personal data. The **potential misuse of data**, including inappropriate data merging or public disclosure of confidential information, poses serious ethical and legal challenges. Ensuring the protection of sensitive information, such as interest declarations, is paramount. Institutions should therefore navigate data protection, and ethical standards and regulations while implementing AI. Moreover, integrity bodies should adopt robust data governance frameworks and implement comprehensive data protection measures to mitigate privacy risks, including conducting regular audits, employing advanced encryption techniques, and fostering a culture of ethical data usage within their organisations (see Chapter 4, section on "Creating a strong data foundation").

Implementation challenges

- Skills gaps
- Lack of high-quality data and the ability to share it
- Lack of comprehensive guidance and practical frameworks
- Risk aversion
- High costs of AI adoption and scaling

OECD (2024[89]) work has found that **a shortage of relevant skills and experience** is one of the biggest challenges in the adoption of AI in public integrity and oversight. This encompasses technical skills (both advanced digital expertise and the upskilling of general staff), as well as change management, data governance and leadership skills.

Access to **comprehensive, high-quality data is a pre-requisite** for the effective use of AI by any institution, yet integrity actors often lack data autonomy**.** While some integrity actors "own" the data they use (for example, interest declaration databases), many – including ACAs and SAIs – often rely on externally generated data over which they have little or no control. They depend on government rules and regulations, historical practices and memorandums of understanding (MoUs) with other institutions to acquire the data necessary for their activities. This means that integrity bodies often face delays in accessing the necessary data, hindering real-time analysis and timely decision-making.

Inconsistent and low-quality data can also severely impede the development and effectiveness of AI tools, making it challenging to create reliable and robust systems. Integrity bodies often rely on data from

external sources, making the availability and accuracy of this data critical for successful implementation. The quality of externally sourced data is frequently unknown or uncontrolled until it is received, potentially leading to inaccuracies and inconsistencies that can limit or complicate AI development and system training. As quality data are foundational the effective development and use of AI, regulatory frameworks need to facilitate data sharing and collaboration between government agencies and across sectors, as appropriate and consistent with data protection and privacy rules, and strategies and systems should place to ensure data quality and completeness.

Data interoperability is also a barrier when it comes to sharing or combining data. The ability to cross-reference and analyse relationships within data is central to uncovering issues such as conflicts of interests. Yet currently, the maturity of most governments' data platforms means that cross checking information is extremely difficult, if not impossible. Regulatory constraints – for security or protection reasons – may further impede the ability to integrate different data sources, thus limiting the potential for effective and comprehensive oversight.

Perhaps because of the nature of the work, integrity institutions are often risk averse. Many integrity actors want to integrate advanced technologies but are unsure where to begin or how to proceed. This apprehension is made stronger by **the fear of making mistakes in the adoption process**. Government guidelines frequently focus on cautionary advice, emphasising what not to do, rather than providing concrete, actionable steps for effective AI implementation. This lack of practical, implementable advice leaves many integrity bodies without a clear path forward, stalling progress and innovation. Comprehensive **guidance and practical frameworks** on the trustworthy use of AI for integrity actors, which are often lacking, actors could help.

Finally, **the cost of implementing and maintaining AI tools poses a significant challenge**, particularly for institutions with limited budgets and technical expertise. The complexity of ensuring robust cyber risk mitigation measures further complicates this landscape. In response, collaboration across institutions becomes essential. By using centralised resources, such as government-led AI sandboxes or shared tools, integrity bodies can reduce costs while accessing advanced technologies.

Untapped potential and way forward

Looking ahead, adaptive learning systems could continually refine their predictions based on new data they can play a significant role in updating risk assessments dynamically, which will allow integrity bodies to stay ahead of potential evolving threats. AI — particularly its ability to analyse and recognise patterns in non-numeric data — can also be the source of future improvements in anti-corruption. By combining it with sensor and image technology (satellites, drones, aircrafts capturing images or other data such as thermal readings or radar signals), AI creates new opportunities to monitor and analyse patterns on a large-scale to understand corruption-related activities and dynamics (Zinnbauer, 2025[105]).

Further down the line, agent-based modelling (ABM) by modelling entities (officials, businesses, etc.) and their interactions, could simulate how corruption and influence evolve in different conditions (U4, 2025[106]). Doing so could help test the impact of anti-corruption or integrity policies, laws and strategies before implementation.

AI-driven network analysis could also help more efficiently and comprehensively map the relationships between lobbyist, politicians, corporations and other actors, revealing influence patterns, conflicts of interest and revolving-doors dynamics. While integrating AI technologies into public integrity and anti-corruption efforts offers great potential, ensuring their effective and ethical use requires careful attention to key policy considerations that promote trustworthy adoption. Recent OECD (2024[89]) work provided three pieces of advice for integrity actors in exploring the use of AI:

- Start by incorporating generative AI into low-risk areas and processes. This approach can help build capacity in areas where mistakes are less costly — either financially or from a compliance perspective — before they scale to riskier and more analytical tasks, including those that require more financial resources.

- Consider the IT requirements for both piloting and scaling. Consider what computational and storage resources, data storage and data management capabilities may be needed and make sure early decisions do not overly complicate future plans.

- Consider internally generated or open data to demonstrate value and establish quick wins. This approach is low-cost and helps demonstrate use cases faster.

As government continue their exploration of AI, they should establish robust transparency and accountability mechanisms to ensure AI-driven decisions are not only accurate but also appropriately interpreted and documented. For instance, while AI systems can provide valuable insights by processing vast amounts of data, and future systems could be increasingly autonomous, the ultimate authority over sensitive decisions need to remain with human operators to preserve fairness, professional judgement (or scepticism) and ultimately maintain public trust. Public integrity bodies should seek to ensure that: the reasoning behind decisions is understandable both to experts and lay audiences; there are recourse and challenge mechanisms in place; and, to the extent possible, AI processes are explainable. AI-generated evidence needs to meet documentation requirements to provide an appropriate audit trail, including complying with current standards to demonstrate how evidence was obtained and assessed.

Knowledge exchange is also crucial for the successful implementation and scaling of AI technologies in public integrity roles. Institutions should participate in and create platforms that enable the sharing of best practices, strategies, lessons learned and even models, code and data. Sharing will help build a more comprehensive understanding of AI's effectiveness in integrity roles and drive the development of best practices that can be widely adopted. Initiatives like the OECD Anti-Corruption in Government division's Tech and Analytics Community of Practice or the Tech Connect for Integrity initiative serve as examples.[25] These forums bring together public and private sector practitioners to discuss challenges, explore innovative solutions, and develop collective strategies to enhance the adoption of AI and other technologies. This can allow for better flow of information, pooling of resources, rapid learning and faster development and implementation of methodologies and solutions.

Building on existing guidance on the ethical and responsible use of AI, governments should go beyond broad directives and provide more detailed, actionable guidance and frameworks for implementing AI technologies. This may include comprehensive step-by-step instructions, case studies and templates that institutions can tailor to their specific needs. Such guidance could also provide advice on identifying and mitigating bias in data, utilising explainable AI techniques, building transparency into decision-making processes and how to incorporate redress mechanisms.

A unified and interoperable data foundation is essential for the effective use of AI. Governments need to streamline their data operations to ensure that databases and systems can seamlessly communicate and integrate. This involves creating standardised data formats, establishing robust data governance frameworks and investing in infrastructure that supports data interoperability. By promoting data standardisation and integration, as well as policies for safe data sharing and crossing, governments can enhance the ability of AI systems to analyse data comprehensively, identify patterns and provide actionable insights to support public integrity.

AI in policy evaluation

Public policy evaluation provides crucial evidence to help policymakers understand what works, for whom and under what circumstances (OECD, 2025[107]; 2020[108]). AI can play a significant role in supporting policy evaluation. The exponential rise in globally generated data coupled with the rapid development of new technologies lowering storage and computational costs, is driving innovation in techniques that can instantly capture, analyse and visualise these vast data repositories, enriching evaluations (Petersson et al., 2017[109]) (Rinaldi, Giuffrida and Negrete, 2017[110]). This can have important consequences on what and how evaluators evaluate.

AI has the potential to accelerate and automatise essential tasks, such as data collection and analysis, and support evaluators in different managerial tasks, by accessing a broad set of internal and external data, or synthesising results. AI can also support ex ante evaluations by building predictive systems and simulations that help policymakers anticipate potential impacts before implementation (Bénassy-Quéré, 2022[111]). However, its use within government for policy evaluation has been limited and has progressed slower than in other functions discussed in this chapter.

Current state of play

Supporting evaluation design and implementation

AI can help policy evaluators to process a large amount of content that can be useful to design and implement evaluations, whether ex ante or ex post. This ranges from more sophisticated and structured approaches, such as supporting synthesis of existing evidence, to more basic functions, such as providing plain summaries of previous evaluations.

Evidence synthesis is particularly useful to inform the development of ex ante or ex post evaluation. These methods typically involve combining results from different studies that investigate the same topic to have a comprehensive understanding of the subject. In the field of evidence synthesis, text mining and other types of AI tools have been applied to literature searching, screening full text, data extraction and analysis for more than 10 years. Researchers in the field have mapped a variety of tools that can support the different stages of evidence synthesis, particularly for systematic reviews and make them more effective. Examples of tools include Rayyan, for screening titles and abstracts, and Robot Reviewer, for assessment of certain risks (Khalil, Ameen and Zarnegar, 2022[112]). Guidance and recommendations were developed to ensure the responsible use of AI in Evidence Synthesis (RAISE), showing the growing interest of applying these tools in the field (Thomas et al., 2024[113]). The benefits of using AI for evidence synthesis rely on the ability to access a vast amount of literature and to process the information faster than alternative approaches. Indeed, the average time to conduct an evidence synthesis is 15 months, but the use of AI can radically reduce some of the steps (Blaizot et al., 2022[114]). For example, a risk of skewed data assessment of 30 randomised control trial (RCT) articles can be accurately conducted using an LLM in an average of 53 seconds. That would have taken orders of magnitude more time for human; a recent study provided an estimate of around 28 minutes for each study in the systematic review using a tool for randomised trials (RoB 2), which is now considered the gold standard (Minozzi et al., 2020[115]; Lai et al., 2024[116]; Odell, 2024[117]). Similarly, World Bank evaluators, using text mining and AI, have been able to double the size of the evidence base they use to make certain programme decisions. That likely would

have been impossible using traditional portfolio identification approaches, laying the groundwork for the relevance and effectiveness assessments in the portfolio analysis they conducted (Bohni Nielsen, Mazzeo Rinaldi and Petersson, 2024[118]).

Supporting analytics

AI can support processing large quantities of data and long texts, such as documentation reports and interviews, to identify patterns, which can be particularly useful for evaluations. By using NLP-enabled text mining, AI can help evaluators comprehend conclusions and provide fine-grained assessments that would otherwise be time-demanding (Næss et al., 2025[119]).

Quantitative text analysis applications for programme evaluation are promising, with AI enabling faster analysis of a large number of documents compared to classical text analysis tools (Gatto and Bundi, 2025[120]). For instance, analysing responses to open-ended questions required to reconstruct a programme theory could involve different time-consuming methods, such as identifying statements that have a specific form and reformulating such statements into conditional "if–then" propositions (Leeuw, 2003[121]), which can be facilitated by novel techniques, such as topic models (Gatto and Bundi, 2025[120]). While conventional text analysis can contribute to evaluation in several empirical and conceptual ways — such as by measuring stakeholder preferences or identifying underlying programme theories — AI-enabled quantitative text analysis can create even more opportunities. For example, the World Bank used unsupervised ML to analyse 392 project reports across 64 aid-receiving countries, successfully identifying novel and insightful factors influencing project success and failure (Franzen et al., 2022[122]). AI text analysis is particularly helpful in identifying underlying themes in existing policy programmes or reports by mapping key concepts embedded in relevant documentation, a technique known as a topic modelling (Cintron and Montrosse-Moorhead, 2021[123]). Some of these methods can be used to understand positions of different stakeholders providing interesting insights for policy evaluation. In some countries, relevant actors, such as SAIs, are adopting quantitative text analysis techniques to conduct performance audits that share several characteristics with policy evaluations. In Norway, for example, text mining and ML applications were used to perform a performance audit on police handling of cybercrime (Box 5.31). In the domain of environmental policy, the OECD partnered with leading research institutions to use AI to conduct the first comprehensive global evaluation of environmental policy measures, analysing over 1 500 policy interventions across four sectors from 1998 to 2022 in 41 countries across six continents (Box 5.32).

Beyond text analysis techniques, AI also has potential to enhance causal inference in policy evaluation, to support quasi experimental designs that rely on probabilistic models, helping to generate some of the missing data that are needed for sophisticated non-parametric methods and instrumental variables to simulate various impact scenarios. This can help move the frontiers of research and enable way more powerful evaluations without the need for costly randomised control trials approaches (Miller, 2020[124]).

Box 5.31. Norway uses text mining and ML in police audits

In 2018, Norway's Office of the Auditor General established an Innovation Lab aimed at integrating data science into auditing. The lab included data scientists with expertise in areas like performance evaluation/auditing, coding and ML, to support auditors by automating tasks, assisting with data acquisition (such as web scraping) and analysis (e.g. text mining), as well as by providing internal training in coding and analytical tools.

In 2021, the Office of the Auditor General of Norway conducted a performance audit of the Norwegian national police efforts at combatting cybercrime. The goal of the audit was to examine whether the Norwegian police has improved its ability to combat cybercrime. During the audit, the collaboration between data scientists and auditors was key in ensuring a skilful use of novel methods in such audits,

providing contributions in their respective competences. While the auditors dealt with several tasks, such as research design, data collection and analysis, as well as writing, data scientists played a key role in applying advanced methods of text mining and ML to classify criminal cases. More specifically, during the audit, 1 000 coded cases were used to train a ML system that applied text mining techniques to extract the text from 334 544 cases across different types of crime (e.g. illicit gain/theft, traffic, violence, financial). The task of classifying criminal cases as cybercrime or non-cybercrime was crucial for assessing the police's understanding and handling of such crimes. The data previously held by the Norwegian police was unreliable and lacked insight into the extent of investigations and case resolutions.

Source: (Næss et al., 2025[119]), https://www.adb.org/sites/default/files/publication/928976/governance-brief-052-digital-transformation-tax-administration-rok.pdf.

Box 5.32. Evaluating the effectiveness of environmental policies with AI

Understanding which policies effectively reduce emissions is a pressing challenge for governments worldwide. In partnership with the Potsdam Institute for Climate Impact Research (PIK), the University of Oxford and the University of Victoria, the OECD has contributed to a study that uses AI to evaluate the effectiveness of such policies across 41 countries. This research, published in *Science*, takes stock of over 1 500 policy interventions spanning 1998 to 2022, offering unprecedented insights into what works in the fight against climate change.

At the heart of the study is the OECD's Climate Actions and Policies Measurement Framework (CAPMF). Using AI, the study identified 63 policies that have led to significant emission reductions. This AI-driven approach not only enhances the ability to detect successful policy measures, but also uncovers patterns and policy mixes that might otherwise go unnoticed. One key finding highlights that policy combinations — rather than stand-alone measures — are most effective in reducing emissions, reinforcing the importance of strategic policy design.

Source: https://www.oecd.org/en/blogs/2025/01/what-works-groundbreaking-evaluation-on-the-effectiveness-of-climate-policies.html.

Supporting management and communication

Evaluation managers can benefit from a variety of AI-based tools developed to facilitate support activities, such as administrative processes, drafting, translating and searching tools. When looking at planning and management, government evaluations are often outsourced to external evaluators, increasing their managerial complexity. The AI-based features of project management tools, such as monday.com or Asana, help evaluators optimise resources and timelines, for example, by recommending optimal staffing and scheduling based on prior evaluations, or ensuring resources are allocated where they are most needed. Both platforms use AI to provide automation and insights, such as workflow automation, predictive tasking or assistance through AI assistants. In addition, generative AI can also facilitate writing terms of reference or similar managerial tasks (Jacob, 2025[125]).

Tools using LLMs can help to improve the communication of evaluation results; they quickly summarise long reports into shorter products that can be shared with decision-makers or the public. The European Commission, for example, has developed an LLM-driven tool that can produce summaries and policy briefs from documents produced in different languages (Box 5.33).

Finally, AI tools can be useful to develop digital evaluation repositories and knowledge management tools. Several OECD countries have developed evaluation repositories, which ensure that all conducted evaluations are easy to find, or platforms to describe and align evaluation efforts across government. One example is Norway's Kudos platform.[26] While the manual development of such repositories can be time consuming, LLMs can automate much of the work and allow for more precise searches across a large volume of reports. In France, the General inspectorate of Finance (IGF) is currently using LLMs to develop an internal retrieval augmented generation (RAG) (Box 5.34) tool called Fragments, which collects reports by the IGF and French Court of Audit since 2006 and allows for precise searches through these documents.

Box 5.33. European Commission's eSummary tools to support public administrations

The European Commission's Directorate-General for Translation has developed a series of AI-based services that can support policy analysis including policy evaluation. For example, eSummary is an AI-based service that can perform a quick overview of a submitted text and send a shortened version back; it uses AI algorithms to choose where the emphasis lies in the document and provide a relevant synthesis. eSummary is connected to another AI-based translation tool (eTranslation), allowing it to create text in all the EU languages. The tool is accessible to a variety of actors across the European Union, including public administrations in member States.

Source: https://language-tools.ec.europa.eu.

Box 5.34. Retrieval-Augmented Generation for LLMs

Retrieval-Augmented Generation (RAG) is a technique developed to improve how LLMs, such as those behind advanced chatbots and virtual assistants, handle information. For different reasons, including reliance on old data, LLMs can provide incorrect answers, and it can be difficult to understand how they arrived at a particular response. RAG allows LLMs to access additional data sources that can keep information current, which is particularly useful when applied to specialised domains or knowledge areas. For government actors, RAG can be an effective means of fencing in their internal data sources, while improving the accuracy, relevancy and trustworthiness of a model's output.

RAG begins with identifying pertinent documentation and extracting vital text from it. Then, it breaks this text down into smaller parts and transforms these parts into a format (i.e. embeddings) that the model can understand and store efficiently. These pieces of information are kept in a special database (i.e. vector databases). When someone asks the model a question, it can look through this database to find up-to-date and accurate information to add to what it already knows before giving an answer.

For situations where it is critical for a model to provide facts that are current and accurate, such as when dealing with confidential information or needing to keep a clear record of data sources, the United Kingdom's AI Playbook for the UK Government (Box 4.2) recommends using RAG. This approach can help to ensure that the model's answers are based on reliable data, making it particularly valuable for organisations focused on maintaining high levels of accuracy and accountability.

Source: (Ugale and Hall, 2024[89]).

Evidence of impact

As the use of AI in policy evaluation is still at early stages, the impact of AI on the practice of policy evaluation is still modest and difficult to measure. A recent study conducted on 758 Boston Consulting Group (BCG) consultants on complex, realistic knowledge-intensive tasks (such as policy evaluation), showed that each one of a set of 18 realistic consulting tasks within AI's known capabilities, consultants using AI were significantly more productive. Compared to a control group, on average, they completed 12% more tasks, 25% faster, with 40% higher quality (Dell'Acqua et al., 2023[126]). However, a different effect was observed for tasks "outside the current capability of AI", showing fewer mistakes from the group not using AI. Other studies on the impact of AI on evidence synthesis show promising results of using LLMs to conduct some elements of systematic reviews, like risk of bias assessments, with agreement in judgment between humans and AI systems ranging from 41% for overall judgement to 71% for outcome measurement (Eisele et al., 2024[127]). At the same time, the study also highlights that AI judgment cannot yet replace human assessment.

Managing risks and challenges

While there is very little research on the risks and challenges of using AI in policy evaluation (Jacob, 2025[125]), that research — along with OECD work with governments and analysis of individual use cases — have identified several associated risks and implementation challenges for AI use in the field.

Associated risks

- Inadequate or skewed data in AI systems
- Automation bias
- Lack of transparency and explainability

AI's use during evaluation processes can reinforce some errors that can emerge at multiple points within the AI pipeline. The first risk of using of AI in policy evaluation originates from **training algorithms with skewed or incomplete data. This** can produce systems that generate erroneous predictions or that reinforce or exacerbate pre-existing outlooks (Jacob, 2025[125]). While this risk is does not only concern policy evaluation — and can apply to other uses of AI in policymaking — its potential risk in the field are significant considering the impact on maintaining or discontinuing potentially positive or harmful interventions (Marcucci and Verhulst, 2025[128]). For this reason, it is essential to mitigate this risk as much as possible by making sure that the data used to train these systems is of good quality and is representative.

Many people perceive AI systems and the decisions they make to be neutral and impartial, leading them to accept results without scrutiny, despite the possibility of inaccuracies. This tendency for human operators to over-rely on automation is known as "**automation bias**" (Horowitz, 2023[129]; Alon-Barkat and Busuioc, 2022[130]). Over-automation could reduce the role of human judgment and potentially oversimplifying complex social and economic assumptions. This may cause evaluators to accept recommendations proposed by AI without fully scrutinising the underlying assumptions or data.

Furthermore, the **lack of transparency of certain AI tools** can further complicate policymakers' tasks to understand and justify AI-driven insights. These can be particularly problematic in policymaking, where an objective approach is needed to address proportionately different populations and justify the decisions made.

Implementation challenges

- Lack of high-quality data and the ability to share it
- Skills gaps

Just as governments face a number of challenges in promoting policy evaluation, they also face specific challenges when seeking to use AI in this field (OECD/UNESCO, 2024[131]). **Inadequate data governance** has long been a challenge for evaluators, limiting the capacity of governments to generate the data that is necessary to produce evidence and evaluation (OECD, 2020[108]). This issue persists even though in adhering to the OECD (2023[132]) Recommendation on Public Policy Evaluation, all member countries committed to ensuring the availability of high quality, timely, accessible, disaggregated and reusable results, performance and administrative data for policy evaluation.

The implementation of AI requires **digital and numeracy skills**. While not all members of an evaluation team need to be experts, it is important to ensure that all evaluators understand how AI can support policy evaluation. Evaluators should therefore be properly supported by digital and data science teams, which are currently being put in place, such as at France's Inspectorate General of Finance (IGF). Evidence shows that policy evaluators have been slower to adapt to the new developments at a general level, even if some countries have advanced practices. There is often limited development training in big data analytics and AI for evaluators. This is also the case for evaluations inside government, which often suffer from limited internal analytical capacities and technical skills, hindering the development of quality evaluations (OECD, 2020[108]).

The OECD (2025[107]) *Implementation Toolkit for the OECD Recommendation on Public Policy Evaluation* provides information on how governments can assess their current policy evaluation capacities, identifies best practices and use cases from around the world, and illustrated potential practical solutions and tools for policymakers and evaluators.

Untapped potential and way forward

In the field of policy evaluation, AI can perform some tasks, allowing government analysts to use a broader range of evidence and process it faster. While some initial applications of AI were identified in evaluation design, analysis, and evaluation communication and management, the use of AI in policy evaluation remains limited. For this reason, there are areas where AI has the potential for significant impact on policy evaluation in the future.

First, to support evaluation design, chatbots could help evaluators build their knowledge in specific fields. Indeed, if well prompted, chatbots can perform several activities that can support learning. As some initial examples of evaluation design show, they can also support creative thinking and be used as useful tools for brainstorming (Ferretti, 2023[133]). Even if these tools do not generate new evidence, they can provide new insights helpful for initial stages of an evaluation process. Recently, for instance, ChatGPT's Deep Research attempts to automate a large part of the process of evidence review and synthesis. Using chain of thought (CoT) reasoning, tools such as Deep Research break complex research questions into smaller, comprehensible sub-questions that it answers in sequence. The approach enables the system to prepare a detailed report based on its review of the available evidence. Such CoT techniques have the potential to automate a large part of the evidence review and synthesis process. This may enable researchers who would previously develop a few reviews from scratch to instead automate, quality assure and build upon dozens of AI-generated research reviews.

Second, from an analytical perspective, there is a strong potential for AI to be further used to conduct more ambitious ex ante and ex post evaluations, using a broader range of data and assessing impact through quasi-experimental methods. For instance, AI-driven behavioural forecasting can analyse large quantities

of historical data and observed behaviour to identify patterns, anticipate decisions and optimise user experiences by integrating contextual variables and external stimuli. ML tools can be used for counterfactual prediction in cases where a control group is missing. This can be used, for example, in the case of carbon pricing assessments, where policy evaluators lack an ex-post perspective. One study proposes a policy evaluation approach using ML tools and economic theory for counterfactual prediction to analyse the costs and emissions impacts of the UK CPS, "a carbon tax levied on all fossil-fired power plants" (Abrell, Kosch and Rausch, 2022[134]).

Finally, in the longer-term, AI has the potential to change the approach to policymaking from a policy cycle and allow evaluations to feed into decision-making at multiple stages. Because AI makes evaluations quicker and to some extent less costly, academics suggest the possibility of shifting from a system where evaluations often arrive too late for decision-making to an approach where evaluative evidence is available to shape, adjust and redesign policies almost in real time. This is referred as Dynamic Public Policy Cycle (Jacob, 2025[125]). As countries around the world have faced a series of crises in recent years, it is essential that governments have access to evaluative evidence in key decision-making stages. Rapid evaluations are developed to inform urgent decision-making and have been efficiently used for this purpose, for example, in Australia (Better Evaluation Knowledge, 2022[135]). While these rapid evaluations now rely mainly on qualitative data, AI could play an important role in making these evaluations more robust and common in the future.

However, for AI to effectively support evaluation, governments need to invest in civil servants' skills and developing a strong data infrastructure. Stronger international collaboration can also enhance AI's potential in policy evaluation. Evaluators need a good understanding of AI's potential benefits, risks and limitations to make informed decisions on when and how to use it. For this reason, governments need to invest in training courses for evaluators to ensure they understand the different tools available to them. Trainings have been developed across OECD governments (see Chapter 4, section on "Fostering skills and talent"). However, these are mainly on the use of AI in government and are not tailored to the field of evaluation. In addition to trainings, it is important to support experimentation with AI and learn-by-doing. Developing a network across line ministries to exchange relevant AI applications can be a good way to support AI uptake in different evaluation tasks. Some incubators are currently being developed but a stronger focus on evaluation is needed.

As is the case for other policy areas, governments should invest in relevant data infrastructures and data sharing that is safe and secure (see Chapter 4, sections on "Creating a strong data foundation" and "Building out digital infrastructure" for a detailed discussion). Some government organisations, such as the Australian Centre for Evaluation, have developed guides to facilitate data discovery and access to support evaluation activities (ACE, 2025[136]). Some OECD countries have developed ways in which different datasets can be linked and accessed in secure environment to ensure policy analysis. In Denmark, for example, Statistics Denmark (2025[137]) facilitates the use of these micro-level databases for research purposes for approved analysts, universities, research organisations or ministries. In the Netherlands, the government launched the Data Agenda Government, outlining plans to improve the management of personal data, open data and big data, leveraging analysis and integration for informed policymaking and addressing societal challenges (Netherlands Ministry of the Interior and Kingdom Relations, 2019[138]).

Finally, AI has potential in evidence synthesis. There is a broader calling for stronger collaboration on evidence generation across countries, following strategic initiatives supported by countries such as the United Kingdom and Australia (Halpern and Maru, 2024[139]). This agenda recognises a need for faster, reliable synthesis at the international level, given that AI is already helping to reduce timelines for evidence production. This could help to fill some of the existing gaps faster.

AI in civic participation and open government

Stakeholder participation refers to "all the ways in which citizens and stakeholders can be involved in the policy cycle and in service design and delivery" (OECD, 2017[140]). Civic participation, which focuses on engagement with the public, can yield benefits for governments and citizens alike. It not only builds trust in government, but it also reinforces democracy by providing meaningful opportunities for citizens to have a say in how government designs and implements policies and services (OECD, 2023[141]; 2024[142]; 2022[143]).With regard to AI policies and products, engagement early in the technology development cycle enriches the understanding of issues, helps align technological innovation with societal needs, and influences how well potential harms are mitigated (OECD, 2024[144]; UK Government Office for Science, 2023[145]).

Governments are using digital technologies to expand civic participation opportunities by enabling new forms and channels of interaction between citizens and public institutions. This contributes to the growing trend of Civic Tech: the use of digital technologies to reinforce democracy by enabling the public to be informed, participate in decision and policymaking, and increase governments' responsiveness and accountability (OECD, 2024[146]). In particular, AI tools have relevant applications in civic participation and more broadly in open government policies and practices, including access to information, government communication and protection of the civic space.

Current state of play

Improving transparency and information access

AI's ability to sort, filter and summarise vast amounts of information could lower barriers to disclosure, making it easier for governments to be transparent to the public (OECD, 2024[29]). For example, the US government is testing an AI-powered tool to review classified documents for future disclosure under the Freedom of Information Act (US FOIA, 2023[147]; MITRE, 2023[148]). AI can also help citizens understand complex governmental processes and create opportunities for broader engagement and scrutiny from citizens and civil society organisations (CSOs) (Araszkiewicz and Rodríguez-Doncel, 2019[149]).

Governments can offer AI-enabled tools to the public directly to help people navigate and make sense of extensive amounts of data and information, thereby providing citizens with tailor-made and accessible access to public information. For example, the District of Columbia (US) has released a beta version of the platform DC Compass,[27] which uses AI to respond to users queries and generate maps based on the city's open data. In Greece, the opencouncil.gr platform uses AI to automatically transcribe local council meetings and generate summaries, social media content and personalised neighbourhood updates via messaging apps, making local governance more accessible and understandable for citizens.[28]Another example from Greece is the platform DidaktorikaAI launched by the National Documentation Centre (EKT), improves the accessibility of academic and scientific knowledge for policymakers and the broader society through an AI-powered online library gathering more than 50 000 publications (National Documentation Center (EKT), 2025[150]). Furthermore, the European Parliament provides an AI tool to help citizens access its archive by asking questions in natural language (Box 5.35).

> **Box 5.35. Democratising access to the European Parliament's archives**
>
> The European Parliament Archives have embraced generative AI to transform access to their extensive historical records. Developed by the Archives Unit, the innovative Archibot platform integrates LLMs with RAG techniques to enable users to interact directly with digitalised documents. Citizens, researchers and policymakers can now make natural language queries to retrieve information on historical events, key figures and institutional developments from an index of over 100 000 documents dating back to 1952. The tool employs RAG methods to identify relevant documents based on users' queries and synthesizes the extracted information into concise, easily comprehensible responses. Accessible via the European Parliament's archive website, the platform supports automated translation, thereby extending its reach to users in 55 languages despite the predominantly French content.
>
> The project, which reduces document search and analysis time by approximately 80%, significantly improves the efficiency of historical research and data analysis. In addition to facilitating rapid access to legislative and policy information, Archibot supports the creation of comprehensive reports and analytical outputs, serving as a valuable resource for educators, researchers and policymakers. This digital transformation of archival access not only addresses longstanding challenges in navigating vast records collections but also enhances transparency and public engagement within the European Union.
>
> Source: https://historicalarchives.europarl.europa.eu/en/sites/historicalarchive/home/cultural-heritage-collections/news/ai-dashboard.html.

Enabled in part by improved transparency and access to information, AI tools present opportunities to promote civic space (OECD, 2023[151]). Civic space is a cornerstone of functioning democracies. It consists of the set of legal, policy, institutional and practical conditions non-governmental actors need to access information, express themselves, associate, organise and participate in public life (OECD, 2022[152]). For example, in Israel, the platform Kol Zchut (All rights) — a comprehensive repository of residents' rights developed by the homonymous CSO along with the country's Ministry of Justice and National Digital Agency — has recently integrated an AI chatbot to allow users to formulate their queries in Hebrew, Arabic and Russian (OECD OPSI, 2019[153]). AI experts have identified empowered citizens, CSOs and social partners (e.g. trade unions) as one the most critical potential benefits of AI use, calling on governments to take more action to empower these actors (OECD, 2024[29]).

AI can also expand access to information by helping people with disabilities that may face barriers to their access to information. For example, citizens affected by visual or cognitive impairments can benefit from speech-to-text or image-to-speech conversion applied to government websites and portals, as well as from AI-powered accessible interfaces with adaptive texts, fonts, colours and contrasts (Welker, 2023[154]). The United Kingdom built its gov.uk website to be fully compatible with technologies that use AI to improve its accessibility. AI can also enhance access to information by overcoming the language divide, including on government websites through automatic translation of content, which has become mainstream. This can enhance the ability to engage in participatory processes by fostering dialogue among participants speaking different languages (European Parliament, 2022[155]). India's Bhashini platform, discussed in Chapter 4, serves as another example of seeking to overcome language divides, including through leveraging participatory techniques.

Making participatory processes effective and inclusive

Governments can use AI to help conduct and facilitate public discussions and dialogues. Such dialogues often face a trade-off between breadth (the number and variety of participants) and depth of discussion and consensus-building (Landemore, 2022[156]). While the internet has enabled spaces for many-to-many communication, these spaces are not always designed to foster deliberation and consensus-building

(Weyl, Tang and Plurality, 2022[157]). AI can provide for a welcoming and productive deliberative space by (Landemore, 2022[156]; Ovadya, 2023[158]):

- Extracting meaningful insights from the contribution of many participants, which could be provided under different formats, such as text and voice.
- Detecting of harmful content and supporting human moderators in countering hate speech.
- Facilitating fact-checking by providing data and sourced information in real time.
- Helping to perform live facilitation to favour multiple opinions and to create bridges towards consensus and support deliberation.

AI can also enhance the quality of dialogue. For instance, it can support organisers in drafting accessible and fact-based learning materials and **simplify complex bureaucratic documents into easy-to-read language to enable informed debate**. For instance:

- In the United States, the AI-powered tool MAPLE (Massachusetts Platform for Legislative Engagement) allows citizens to better understand the context and objectives of draft legal texts through AI-generated summaries submitting their inputs and comments.[29]
- In Spain, the participatory platform Decide Madrid, based on the open-source software Consul, experimented with an NLP system to assist citizens in aggregating and developing proposals (Arana-Catania et al., 2021[159]).

Research indicates that the use of **AI-enabled deliberation tools could have a significant impact on decreasing polarisation** (Tsai et al., 2024[160]; OECD, 2020[161]). The examples in Box 5.36 illustrate that such tools are not only beneficial for obtaining understandable insights on a specific topic, but can help bring together the results of the multiple deliberations. They can ensure consistency while preserving the richness of the debates (Landemore, 2022[156]), as well as strengthen connections between the participants and the broader society by making the content and results of discussions more accessible to the public.

Box 5.36. AI as facilitator of deliberative processes

Mass online deliberation with Polis

Polis is an AI-powered open-source software designed to allow large group of people to discuss, collaborate and reach consensus on complex issues. Participants share their opinions on a given question and express their agreement or disagreement on other participants' statements. The algorithm maps participants' responses and clusters opinions, facilitating the identification of positions that are supported by the majority of participants. Polis has been used to facilitate discussion and build consensus in Austria on environmental issues, in Uruguay on a national referendum, in New Zealand to facilitate the development of government policy, in the Philippines to generate municipal policy, in the United States to counteract polarisation at the municipal level, and in Germany to develop a political party's platform.

In Chile, AI helps moderate deliberative processes

In 2020, the Senate of Chile partnered with the Stanford Center for Deliberative Democracy to involve citizens in the online deliberative process "LXS 400 – Chile Delibera" performed through the Stanford Online Deliberation Platform. The platform used an AI tool to keep time and distribute speaking rights equitably, allowing participants to form speaking queues and discuss in small groups with timed agendas. Over 500 citizens representative of the Chilean society discussed in small groups and deliberated in plenary sessions on health and pension reforms.

Source: (The Computational Democracy Project, 2024[162]; Fishkin and al, 2021[163]).

In addition to facilitating engagement, AI can help governments and citizens analyse and make sense of the inputs from participatory and deliberative processes. AI features — such as automation and processing of large amounts of data and text — can facilitate mass, multilingual online deliberation and dialogue. Governments usually face the challenge of analysing large quantities of qualitative inputs received in public consultations, which can be remedied with the help of AI. For example, the United Kingdom's Incubator for Artificial Intelligence (i.AI) has developed the i.AI Consultation Analyser, which aims to expand the government's capacity to analyse citizens' inputs in legal public consultations. The tool is expected to reduce the cost of this activity by GBP 80 million (AI.GOV.UK, 2024[164]).[30]

AI sensemaking tools allow for the identification and evolution of recurrent topics, the automatic clustering of opinions and ideas, the detection of outliers, and the sentiment analysis of the perception of a given issue (Arana-Catania et al., 2021[159]; Berditchevskaia and Baeck, 2020[165]; Schneider and Sanders, 2023[166]). Sensemaking tools can help public servants understand and visualise citizens' priorities and opinions on issues and support them in translating large volumes of inputs into actionable recommendations that are representative of the participants' views. Citizens can also use these tools to understand others' perspectives, inform and craft their own inputs, and better understand the results of participatory and deliberative processes. Examples include France's national *Grand Débat*, or Great Debate (Box 5.37), as well as:

- The Cambridge City Council in the United Kingdom used the AI-powered sensemaking tool of the Go Vocal platform (formerly CitizenLab), which automatically clusters and prioritises citizen opinions to analyse their contributions to its Design Code, saving 50% of the estimated time for manual processing (Fillet, 2024[167]).[31]

- The French Economic, Social and Environmental Committee (CESE) organised a deliberative process (Convention Citoyenne) to inform legislation on the sensitive policy issue of end of life. The AI-powered tool Panoramic, developed by Make.org, allows citizens, journalists and CSOs to access and navigate the discussions and the results of the deliberative process via an interactive interface.[32]

Box 5.37. Improving participation in France's Grand Débat National

In France, the Grand Débat National, launched in early 2019, exemplifies a pioneering effort to democratise public debate through digital innovation. In response to growing social discontent, the French government established this national consultation to gather citizen input on key issues — including employment, taxation and democratic reform — using a dedicated online platform. Artificial intelligence played a central role in managing and interpreting the vast array of contributions.

AI-driven techniques, such as NLP and clustering algorithms, were employed to systematically analyse millions of submissions. By identifying recurring themes and synthesising diverse viewpoints, these tools enabled policymakers to highlight consensus points and emerging trends. This automated analysis transformed raw citizen input into accessible insights, supporting more informed and transparent decision-making.

Source: https://granddebat.fr.

Finally, AI can facilitate innovation in civic participation activities by **generating simulations and visualisations on various scenarios and alternatives**, including in combination with other emerging immersive technologies like augmented or virtual reality (VR).[33] Generative AI tools can help governments to engage with the public on complex projects concerning a physical space (e.g. urban mobility and redesign). These tools can simulate future and alternative scenarios, allowing citizens to visualise the final results and provide feedback for government consideration (OECD, 2023[2]). Generative AI can

significantly reduce the costs of creating visual simulations by automating the design process and generating realistic renderings. For example, the City of Hamburg in Germany, used 3D and AI technologies in the CityScope platform, developed at MIT Media Lab, to engage relevant stakeholders in the decision-making process for 161 viable locations to home refugees (Campagnucci et al., 2025[168]). Similarly, Finland used the tool UrbanistAI to foster discussion on the pedestrianisation of streets (Box 5.38).

Box 5.38. Finland uses generative AI for participatory urban planning

In Finland, the City of Helsinki used generative AI to enhance participatory urban planning through UrbanistAI. This platform integrates AI-powered visualisation tools with co-design features that allow residents and businesses to collaboratively shape urban spaces. During the 2023 Summer Streets initiative, Helsinki used the platform in community workshops, where participants translated their ideas into AI-generated street designs, fostering more inclusive and engaging urban planning processes.

By enabling the real-time visualisation of community proposals, UrbanistAI helped bridge the gap between citizen input and final design decisions. The tool's ability to align proposals with local policies also streamlined the planning process, ensuring regulatory compliance while maintaining creative flexibility. The success of the Helsinki case has inspired broader adoption of the platform, with cities including Tallinn, Pristina and Berlin integrating it into their participatory urban design initiatives.

Source: https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/reimagining-helsinki-participatory-urban-planning-generative-ai.

Supporting effective and tailored public communications

AI-powered analytics software can help public communicators monitor and analyse online debate, audience perceptions and attitudes on a given issue.[34] AI can help them to improve the quality, quantity, accessibility and relevance of information governments deliver to different segments of the population. Moreover, AI tools can enhance the efficiency and effectiveness of public communicators by assisting in brainstorming, automating routine tasks such as media planning, and ensuring consistency in content across communication channels. These capabilities enable communicators to better engage with their audiences, tailor messages to diverse groups, and respond promptly to public concerns, thereby fostering a more informed and engaged citizenry. An example can be found in Box 5.39.

Box 5.39. Innovating public communications with AI in the United Kingdom

In the United Kingdom, the Government Communication Service (GCS) has been developing and testing GCS Assist, an AI-powered conversational tool designed to help government communicators improve the efficiency and effectiveness of their work. The tool integrates professional guidelines, communication frameworks and proprietary audience insights, ensuring that AI-generated outputs align with government standards. Because AI is most effective when used as a collaborative assistant, GCS Assist is tailored to support both experienced and less-experienced professionals by streamlining routine tasks, enhancing creativity and enabling more time for strategic work.

GCS Assist has undergone an iterative development process, beginning with an initial prototype piloted within the Cabinet Office in late 2023. The tool was designed to support government communicators by generating draft texts, plans and strategic ideas while ensuring data security within government

---

systems. The alpha assessment in March 2024 rated the tool's progress as Amber, acknowledging its potential while highlighting areas for further refinement. A phased rollout is now underway, starting with a limited group of users before expanding access more broadly.

Source: https://gcs.civilservice.gov.uk/wp-content/uploads/2024/04/2024-05-GCS-Innovating-with-Impact-Strategy-interactive.pdf.

The production and dissemination of information materials, as well as the assistance and response to citizens' queries, is a fundamental yet resource intense function to enable effective participation. AI can also enhance public communications efforts by curating and tailoring communications materials. Generative AI can, for example, summarise and synthesise information and create content for different audiences in different formats (video, image, text, etc.). Moreover, GenAI can support governments in providing citizens with accessible and evidence-based information, allowing them to better understand the policy issues at stake as well as the role of the government in regulating them. Examples from the Netherlands and Brazil in Box 5.40 illustrate how this can be done in government.

Box 5.40. Practices of AI for information curation in citizen participation

Helping reporting oral deliberations in the Netherlands

The Netherlands House of Representatives implemented the Speech2Write system, which allows its Parliamentary Reporting Office (PRO) to convert voice to text, as well as translate voice into written reports and produce an "almost-ready" report for human finalisation. The system is based on a parliamentary language model that was developed from scratch based on 2000+ hours of audio of proceedings and text files.

Making complex information more accessible in Brazil

The Brazilian Chamber of Deputies is using AI to improve the support given to parliamentarians and to build a more public-friendly legislative institution. Its Ulysses AI Suite uses an NLP system to classify and help citizens navigate the vast amount of information produced by the Chamber, including laws, speeches, data and research documents. Ulysses also helps parliamentarians draft legislative texts and find similar bills already adopted or rejected.

Source: https://www.ipu.org/innovation-tracker/story/artificial-intelligence-innovation-in-parliaments, (Arana-Catania et al., 2021[159]), https://www.camara.leg.br/noticias/548730-camara-lanca-ulysses-robo-digital-que-articula-dados-legislativos.

Virtual assistants such as chatbots can support communications efforts by helping citizens to navigate government information, by responding to their queries in plain language, and by sending personalised notifications and content based on their interests (Androutsopoulou et al., 2019[169]; van Noordt and Misuraca, 2019[170]; Cortés-Cediel, 2023[171]). In some cases, chatbots have been used as a channel for civic participation (see examples in Box 5.41).

Box 5.41. Chatbots supporting participation in Spain and Colombia

In 2021, the city of Madrid (Spain) embedded Clara, an AI-based virtual assistant into its existing participatory platform, Decide Madrid, based on the open-source software Consul. Clara responds to users' questions on the functioning of the platform and of the participatory processes of the city.

In 2022, the city of Bogota (Colombia) deployed Chatico. This virtual assistant supports the interaction between citizens and the local administration by improving the experience of procedures and services and enabling citizen participation. Chatico is available both on the website and via WhatsApp to enhance its accessibility.

Source: https://www.madrid.es/portales/munimadrid/es/Inicio/Actualidad/Noticias/Nace-Clara-el-asistente-virtual-de-Decide-Madrid, https://oecd-opsi.org/innovations/chatico-virtual-agent, https://gobiernoabiertobogota.gov.co.

Strengthening electoral processes

Policy discussions on AI's potential in the electoral process have focused on protecting electoral processes against the risks of AI, namely by empowering society and building resilience, protecting democratic principles, and promoting information integrity. Yet there are also relevant opportunities for AI to strengthen electoral processes by pre-empting fraud and disenfranchisement, and lowering barriers to entry for underfunded candidates (Eisen et al., 2023[172]; OECD, 2024[29]).

AI tools can be used by electoral management bodies (EBM) throughout the electoral period (Juneja, 2024[173]). Before elections, AI tools can be adopted to support voter list management, voter registration, resource allocation planning, forecasting election costs or targeted advertising. For example, in the United States, the Electronic Registration Information Center (ERIC) (2024[174]) uses ML to analyse voter registration and motor vehicle records. This helps states identify duplicate registrations, deceased voters and eligible but unregistered individuals, which improves voter roll accuracy. During elections, AI tools can contribute to campaign and media monitoring, voter verification, polling place monitoring and ballot counting. Finally, AI can support authorities in post-electoral audits by detecting incidents or frauds (Juneja, 2024[173]).

Managing risks and challenges

AI offers significant opportunities to empower citizens and civil society and reshape relationships between governments and their people for the better. However, potential risks persist, which can have direct implications on democratic spaces. While technological advancements may progressively address these issues, policymakers should consider them when adopting AI tools in this area.

Associated risks

- Inadequate or skewed data in AI systems
- Overreliance on AI
- Lack of transparency and explainability
- Resistance: Lack of public understanding about how government uses AI
- Misuse or questionable use of AI, resulting in surveillance and privacy concerns
- Exacerbating social exclusion and digital divides

If AI systems rely upon **inadequate or skewed data**, it could lead to inaccurate or adverse outcomes for some individuals or groups. With regard to civic participation and open government, this could result in inaccurate or imprecise consideration of citizen input into AI-enabled participation processes.

Moreover, deliberative processes are meant not only to harness collective intelligence and enable societal dialogue on specific policy issues, but also to enhance the knowledge of participants as well as their mutual

understanding and empathy (OECD, 2021[175]). **Overreliance on AI** tools to improve the efficiency of deliberative processes might overlook other fundamental aspects of deliberation.

AI's lack of transparency and explainability, especially in complex systems, can decrease trust in AI tools, and when these are used in democratic spaces, it can in turn affect trust in the participatory process and its outcomes. To address this challenge, some governments have launched algorithm registers that disclose detailed and technical information about algorithms, such as, their purpose, design, data inputs, decision-making processes, potential biases, among others (for details and examples, see Chapter 4, section on "Establishing guardrails to guide strategic and responsible AI").

In addition, there are still many unknowns about user perceptions regarding AI in participatory processes and the influence it may have on their contributions and their **willingness to participate**. Experiments are being conducted to assess the reaction of participants to the introduction of AI tools as intermediaries in participatory and deliberative processes (Hadfi et al., 2021[176]; Kim et al., 2021[177]). The use of AI tools to support participants in formulating inputs and contributions should be carefully handled to avoid that these tools compromise the creativity in language and reflection of participants, focusing rather on efficiency and agreement.

Some governments have misused digital technologies for surveillance or even silence populations and digital opposition, thus damaging the civic space (OECD, 2022[152]). The use of AI systems for surveillance purposes, content censorship and improper forms of predictive policing threaten the free exercise of civil freedoms, such as the right to peaceful assembly and freedom of expression. These improper uses risk creating general mistrust among citizens towards the adoption and deployment of AI systems in government functions and public services. The OECD (2022[152]) found that most national AI strategies fail to discuss the impact of AI on the ability to freely exercise rights, although around half propose concrete oversight and redress mechanisms (OECD, 2022[152]).

Finally, **many languages are insufficiently represented in AI systems**, which are mainly trained in English, Spanish, and Mandarin (Peixoto, Canuto and Jordan, 2024[178]) (see Chapter 1, section on "Exacerbating digital divides" for a detailed discussion of the issue). In the context of citizen participation, this means that inputs submitted in other languages might not be processed and valued in the same way, **creating new democratic imbalances** (Romberg and Escher, 2024[179]). For example, to address the language divide and to preserve the Icelandic language, the government of Iceland partnered with OpenAI to train the LLM GPT-4 in Icelandic (Government of Iceland, 2023[180]). Similarly, the University of Turku (Finland) partnered in 2023 with the company SiloAI to build the Poro suit of systems, a family of multilingual open-source LLMs for all European official languages (University of Turku, 2023[181]). Similar efforts exist in for other languages, including Indigenous and endangered ones (OECD, 2023[182]).

Implementation challenges

- Skills gaps
- High costs of AI adoption and scaling

The adoption of AI tools by civil servants in charge of participatory and deliberative processes, as well as by citizen participation practitioners, requires adequate skills. This skillset spans from the identification of relevant uses of AI tools to support participatory and deliberative practices (OECD, 2025[183]), to the selection and purchase or development of AI solutions that meet the defined objectives, to the use of AI tools in the context of participatory and deliberative processes, to the evaluation of the quality of their outputs.

The adoption, either through in-house development or through public procurement, and the use of AI tools to support citizen participation requires a flexible approach including pilot initiatives, testing, and iterative adjustments. Such approach entails significant costs and require an elevated high tolerance to risk.

Challenges related to costs and scaling of AI tools are particularly relevant for citizen participation, ss participatory processes are often organised and implemented by local governments (OECD, 2022[184]).

Untapped potential and way forward

Given the novelty of the use of AI in civil participation and enablement, it is difficult to assess evidence of impact and general lessons learned beyond individual case studies. The use cases showcased in this chapter suggest that AI has the potential to improve the efficiency, reduce the costs and enlarge the scope of participation, ultimately lowering barriers for citizens and CSOs to voice their opinions and provide feedback at various stages of policymaking. The use of AI tools to process citizens' inputs, automate tasks, support the creation of information and communication materials, and provide citizens with 24/7 virtual assistance could significantly reduce some of the costs associated with the design and implementation of relevant processes. These savings, coupled with the growing availability of open-source tools, could enable citizen participation opportunities in contexts with limited resources, such as in smaller cities and less developed regions.

Digital technologies have enabled asynchronous participation and significantly reduced the burden of geographical barriers, allowing for broader collaboration (OECD, 2022[184]). AI has the potential to lower other barriers, namely the language divide, allowing higher levels of citizen participation. Moreover, virtual assistants and chatbots can further improve the accessibility of participatory and deliberative processes by helping participants navigate complex information. Finally, AI tools for moderation, facilitation and sense-making could become essential to addressing the challenge of scale in deliberative processes (Landemore, 2022[156]).

Early efforts could lead to longer-term conceptions of transparency and responsiveness. Digital technologies have long affected how individuals engage with notions such as their so-called "right to know," and voice opinions about complex operations and their outcomes (Margetts, 2011[185]; Song and Lee, 2015[186]). AI technologies also offer novel ways of enhancing service design, delivery and transparency. Individuals may begin to find current operational systems antiquated and difficult to use, and therefore they may voice their opinions about how services should integrate technologies that can enhance access to information (Carrasco et al., 2019[187]). This could help accelerate institutional reforms that aim to improve transparency and accessibility. These may be driven not only by the adoption of AI technologies but also by their capacity to shift public perceptions and expectations of institutional capacity.

Moving forward, governments should explore AI's potential to enhance existing civic participation processes by expanding government capacities of analysing citizens' contributions, enabling mass-deliberation and reducing barriers to participation. Governments should carefully consider the challenges associated with the use of AI in civic participation, namely issues related to increased technological intermediation of interactions between citizens and government, and its impact on public discourse and trust. The adoption and deployment of AI tools for civic participation should follow clear strategies that include transparency and accountability as fundamental principles. Citizens and civil society should be actively involved in the elaboration of such strategies to ensure alignment with democratic principles and enhance trust and acceptability of the process (see Chapter 4, section on "Engagement to shape strategic and responsible AI").

Government services and justice functions

Government services and justice functions are essential to meeting citizens' needs and upholding the rule of law. This includes the design and delivery of public services that are accessible, efficient and responsive, as well as effective law enforcement and disaster risk management to ensure safety and resilience. Justice administration and access to justice are equally critical, providing fair and timely legal processes that protect rights and promote public trust. These areas have the most impact with the public. Governments are using AI in a variety of ways to ensure that they are efficient and effective, and that citizens and residents are safe and treated fairly.